THE ROLE OF CHLOROPLAST CARBONIC ANHYDRASES IN THE DEVELOPMENT OF C3 PHOTOSYNTHESIZING PLANTS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Kevin Matthew Hines

December 2019

THE ROLE OF CHLOROPLAST CARBONIC ANHYDRASES IN THE DEVELOPMENT OF C3 PHOTOSYNTHESIZING PLANTS

Kevin Matthew Hines, Ph.D.

Cornell University 2019

Carbonic anhydrase (CA, EC: 4.2.1.1) is a zinc-bond metalloenzyme that rapidly catalyzes the reversable hydration between carbon dioxide and water with bicarbonate. C_3 photosynthesizing plants contain very highly expressed CAs in their chloroplasts, which can represent over 2% of all soluble protein. However, no single role has been assigned to CA and its activity in the stroma, although they are proposed to have a myriad of duties, from oxidative stress tolerance and plant defense to macromolecule biosynthesis and pH buffering. Previous research exploring the function of chloroplast CA have focused largely on deactivating or deleting βCA1, the most highly expressed CA and the first to be identified in the chloroplast. However, no study has completely removed all CA enzymes from the chloroplast stroma and observed the effects.

Using the model *Nicotiana tabacum* (tobacco), I identified two stromal CAs, βCA1 and βCA5, and produced CRISPR/Cas9 mutants targeting both genes. While the single knockout lines Δ*βca1* and Δ*βca5* had no striking phenotypic differences compared to WT plants, Δ*βca1ca5* leaves developed large necrotic lesions. Leaf development of Δβ*ca1ca5* plants normalized at the high CO_2 concentration of 9000ppm. High CO_2-grown Δβ*ca1ca5* mutants had no measurable defect in photosynthetic capacity when measured at ambient CO_2 Δβ*ca1ca5* seedling germination and development is negatively affected when seedling development

occurs at ambient CO_2. A series of complementation experiments using altered forms of βCA1 were carried out in Δβ*ca1ca5* plants. Constructs expressing full length βCA1 and βCA5 proteins complemented the Δβ*ca1ca5* mutation, but inactivated (ΔZn-βCA1) and cytoplasm-localized (Δ62-βCA1) forms of βCA1 failed to reverse the mutant phenotype. When infected with tobacco mosaic virus (TMV) *Δβca1* and Δβ*ca1ca5* tobacco failed to show the hypersensitive response (HR), while expression of ΔZn-βCA1 restored the response. Thus, stromal CAs play major roles in plant development and defense.

BIOGRAPHICAL SKETCH

Kevin Hines was born in Holyoke, Massachusetts a small city about an hour and a half West of Boston in April of 1988. His mother worked multiple jobs and his father was a printer. He is the youngest of five children.

From a young age, Kevin had always struggled in school. ADHD made attention difficult and undiagnosed autism spectrum disorder caused the noisy classroom environment to become overstimulating. Because of this, Kevin dropped out during his sophomore year of high school when he was 15 years old. After a couple years of working odd jobs and washing dishes, he decided he wanted to earn his high school diploma after all. He graduated from Chicopee High School in 2008 when he was 20 years old and started classes at Holyoke Community College that summer as a pre-nursing major.

Kevin became fascinated with cell biology and switched his major to biology before the fall semester began. He got his first hands-on experience with experimental design during an "Independent Study" course in which he studied an epistatic interaction by crossing two mutants of *Brassica rapa* (Wisconsin fast plants). He graduated from Holyoke Community College with high honors and transferred to UMass Amherst to earn his B.S. Upon Transferring to UMass he was awarded the Massachusetts Community Scholarship, which is given to one person transferring to the UMass from the community college system.

At UMass, Kevin worked in the lab of Dr. Tobias Baskin to study the properties of cellulose synthesis in grasses. In the Spring of 2012, Dr. Baskin encouraged Kevin to apply to the NSF summer REU at the Boyce Thompson Institute located at Cornell's Ithaca campus. Kevin was accepted into the program

and worked in Dr. Maureen Hanson's lab that summer characterizing the encapsulation peptide of CcmN, which is one of the proteins that comprise β-carboxysomes (a bacterial microcompartment found in cyanobacteria). Kevin was captivated by the project, and he knew that he would be applying to Cornell for graduate school to have the chance to work on it even more.

Kevin started his graduate school career in the Fall of 2013 and was awarded the competitive, merit-based SUNY Diversity fellowship for his first-generation college student status. He has since presented his Ph.D. work at conferences in Canada and Japan. His time at Cornell has also uncovered his love of teaching. Besides serving as a teaching assistant for "BioMG 1350:Principles of Cell and Developmental Biology" and "BioMG 2801:Laboratory in Genetics and Genomics," he was also the co-instructor for "WRIT 4860: McNair Scholars Seminar: Writing the Capstone Portfolio," through Cornell's Office of Academic Diversity Initiatives.

In his spare time, he enjoys playing video games and Dungeons and Dragons with his friends. He is also a big fan of his two cats Milo and Squeak. He is engaged to his partner Kristen and they plan to marry in 2020.

Kevin is currently working as a research associate in Dr. Margaret Frank's lab at Cornell and hopes to one day teach at a community college so that he can share his joy of biology to the next generation of scientists from all backgrounds.

For my mother, Mary Menard Dubin.

Thank you for believing in me when I didn't believe in myself. I wouldn't have made it this far without you and your unconditional love and support. You are my hero.

&

For my father, Kevin Hines.

I miss you every day. You were my best friend and my #1 video game nemesis.

&

For my service dog, Cole.

Thank you for saving me; you were a miracle. Good boy.

ACKNOWLEDGMENTS

I would first like to thank my advisor Professor Maureen Hanson, who accepted me into her lab and who has guided me with patience and understanding. Even with her busy schedule, she would always find time for me and her other students when we needed it. It has been a privilege to work with her over the years on this project.

I would also like to thank Myat Lin and Vishal Chaudhari, two post-doctoral fellows with whom I worked with during my time in the Hanson lab. I probably never would have gotten my Western blots to work if it wasn't for their insights and troubleshooting knowledge. Thank you both for not only listening to me complain about my failed experiments, but willingly offering me advise.

Many thanks also go to my committee members Professors Thomas Owens and David Stern. Thank you for all your help and support over the years. Professor Owens is the greatest teacher I've had in my entire college career and his classes were the most informative and interesting courses I've ever taken. I will consider my career a success if I can be half the teacher he is. I would also like to thank Professor Klaas van Wijk, who agreed to proxy my B-exam on such short notice.

This acknowledgment section would be the length of the rest of the dissertation if I thanked my partner and future wife, Kristen Angierski, for everything she has done. You are my best friend and the best cat mom in the whole world. Thank you for always being at my side.

I wouldn’t be here if not for the love and sacrifice of my parents and stepparents. Your support has meant the world to me and has lifted me up even on the hardest of days. Thank you for everything.

Lastly, I would like to thank my students. Whether I mentored you in the lab our taught you all the ins and outs of fly genetics in the classroom, teaching you has been the great joy of my graduate career. I am proud of all of you and can’t wait to see the great things you’re going to accomplish.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ACC	Acetyl-CoA Carboxylase
BiFC	Bimolecular Fluorescence Complementation
BMC	Bacterial Microcompartment
CA	Carbonic Anhydrase
CCM	Carbon Concentration Mechanism
cTP	Chloroplast Transit Peptide
FFA	Free Fatty Acid
HR	Hypersensitive Response
NPQ	Non-photochemical quenching
PEPC	Phosphoenolpyruvate Carboxylase
PSII	Photosystem II
ROS	Reactive Oxygen Species
Rubisco	Ribulose-1,5-Bisphosphate Carboxylase/Oxygenase
SA	Salicylic Acid
SABP	Salicylic Acid-Binding Protein
TMV	Tobacco Mosaic Virus

PREFACE

Chapters 1 and 2 of this dissertation explore the roles of carbonic anhydrase (CA) in the chloroplasts of C_3 photosynthesizing plants through reviewing and performing CA-mutant studies. The work in Chapter 2 specifically is inspired by the focus of Chapter 3: engineering functional β-carboxysomes (from the cyanobacteria *Synechococcus elongatus* PCC7942) in the chloroplasts of C_3 plants in order to improve photosynthesis.

Chapter 3 contains the work performed early during my graduate career. It is a set of bimolecular fluorescence complementation (BiFC) assays which aimed to determine whether the protein-protein interactions essential for the self-assembly of the carboxysome bacterial microcompartment could occur in the environment of the chloroplast stroma. This and other research[1,2] suggests that it is indeed a possibility that fully functional carboxysomes, along with other elements of the cyanobacteria carbon concentration mechanism (CCM), could be engineered in chloroplasts. Part of this process, as examined in Chapter 1, involves removing CA enzymes from the chloroplast stroma.

The protein shell of the carboxysome is selective for the negative charged bicarbonate, which it then converted into carbon dioxide inside the shell in the vicinity of the large Rubisco (Ribulose-1,5-Bisphosphate

[1] M. T. Lin et al., beta-Carboxysomal proteins assemble into highly organized structures in Nicotiana chloroplasts. *Plant J* 79, 1-12 (2014)
[2] M. R. Hanson, et al., Towards engineering carboxysomes into C3 plants. *Plant J* 87, 38-50 (2016)

Carboxylase/Oxygenase) core, allowing only the carboxylase activity of Rubisco to occur. While bicarbonate ions are abundant in the cytoplasm of *S. elongatus*, you will not find any free CA enzymes. CA is only found inside the shell of the carboxysome catalyzing the critical conversion between bicarbonate and carbon dioxide. Therefore, the overall goal of the project requires stroma CA to be removed from the chloroplasts of C_3 plants in order to engineer functional carboxysomes and other components of the cyanobacterial CCM.

Chapter 2 catalogs the work performed using the CRISPR/Cas9 system to remove stroma CAs from the chloroplast of the model plant *Nicotiana tabacum* (tobacco). Most singe CA mutants, as described in Chapter 1, are relatively healthy. CA activity in tobacco can be reduced by 99% and the plants will appear fine, with only minor changes in photosynthesis and other physiological traits.[3] It was within this context that I started this project. The expectation was that even the double mutants, in which all CA was removed from the chloroplast stroma, would have only minor defects. The unique and drastic phenotype I discovered in double CA mutant lines (in which deleting both βCA1 and βCA5 caused necrosis of leaf tissue) became the focus of my research and the main story of this dissertation.

[3] G. D. Price et al., Specific reduction of chloroplast carbonic anhydrase activity by antisense RNA in transgenic tobacco plants has a minor effect on photosynthetic CO2 assimilation. *Planta* 193, 331-340 (1994)

While I departed from the work being done on engineering carboxysomes into higher plants, this work and the carboxysome project are directly linked. There is still work to be done to characterize the strange phenotype of the double CA tobacco mutant, but the mutant itself represents a critical step in the carboxysome project: a C_3 plant with no soluble CA in its chloroplast stroma. These transgenic lines can now be used in the next stages of the project, which include expressing carboxysome genes from the chloroplast genome and introducing a bicarbonate pump into the chloroplast envelope. Importantly, these next steps also have the potential to answer some of the fundamental questions about the double mutant. For example, if the bicarbonate pump can reuse the phenotype, then that result would suggest that the lack of bicarbonate ions in the stroma is a driving factor behind the necrosis phenotype. In sum, Chapter 3 and the other two chapters, while appearing to be about two completely different topics are linked to each other in the grand scheme of the carboxysome project with both applied and basic science inspiring the other.

CHAPTER 1

Working-class hero: The many jobs of carbonic anhydrases in the chloroplasts of C_3 photosynthesizing plants [REVIEW][4]

Abstract

Carbonic anhydrase (CA, EC: 4.2.1.1) is a zinc-bond metalloenzyme that catalyzes the reversable hydration between carbon dioxide and water with bicarbonate. CAs are found in every kingdom of life and consist of seven families (α-, β-, γ-, δ-, ζ-, η-, and ɵ-), each arising independently via convergent evolution. C_3 photosynthesizing plants contain very highly expressed CAs in their chloroplasts, and the enzymes can represent over 2% of soluble protein. However, no single role has been assigned to CA and its activity in the stroma, despite the functions of other CAs (like cytoplasmic CAs in C_4 plants) being well characterized. After decades of research, chloroplast CAs have been assigned a myriad of duties, from oxidative stress tolerance and plant defense to macromolecule biosynthesis and pH buffering. Here we review multiple mutant studies that examine the physiological effects of CA removal and deactivation. Perspectives on what this data means for the future of the field of crop improvement are examined.

[4] Chapter prepared as an invited review for *Biochemical Society Transactions*

Introduction:

The families and mechanisms of carbonic anhydrase

The metalloenzyme carbonic anhydrase (CA, EC: 4.2.1.1) is not only the Protein Data Bank's "Molecule of the Month" for January 2004 (1), but also an essential and ubiquitous family of enzymes found in every kingdom of life (2). CA rapidly catalyzes the reversible hydration between CO_2 (carbon dioxide) and HCO_3^- (bicarbonate) (Equation 1.1). While the interconversion between carbon dioxide and bicarbonate can occur without CA, the uncatalyzed reaction is quite slow (on the order of seconds in alkaline environments). The presence of the enzyme greatly increases the reaction rate to a rapid $k_{cat} = 10^6\ s^{-1}$, which is six orders of magnitude greater than the uncatalyzed rate (3, 4).

Equation 1.1 $$\boldsymbol{CO_2 + H_2O \leftrightarrow H_2CO_3 \leftrightarrow H^+ + HCO_3^-}$$

The ubiquity of the CA enzyme is explained in part by its evolutionary development into seven separate families, α-, β-, γ-, δ-, ζ-, η-, and ɵ-CAs (5), which all arose independently via convergent evolution and, as such, all vary considerably in their amino acid sequences and protein structures (2). Despite the lack of evolutionary history between the separate CA families, they all function in a similar way: (A) The anionic oxygen of OH^--bound-zinc located at

the active site of the enzyme carries out a nucleophilic attack against the carbon atom in CO_2. (B) This produces a zinc-bound bicarbonate ion, which is then released by a water molecule (3). (C) Zinc-bound-OH^- is restored when the water molecule is deprotonated by a neighboring histidine residue. These basic steps are carried out by every known CA family except for ζ-CAs (found in diatoms), which use a cadmium ion instead of zinc (6). The structure of a plant βCA and its zinc-binding pocket are shown in Figure 1.1.

Organisms may encode one or multiple families of CA to carry out an assortment of different physiological functions. For example, humans and other mammals exclusively use αCAs, which efficiently establish and maintain the bicarbonate buffering system in their circulatory system (7), while higher plants use α-, β-, and γ- CAs to carry out a wide variety of different metabolic functions (8). While the roles of many CAs in plants are well defined, the main purpose of chloroplast CAs is not yet thoroughly understood.

CA in higher plants

Inspection of the reaction catalyzed by CA suggests that it could play a role in photosynthesis; for plants which undergo CAM or C_4 photosynthesis, this is true. The carbon concentration mechanism (CCM) utilized by C_4 photosynthesizing plants relies on the conversion of incoming carbon dioxide in the mesophyll cells into bicarbonate (9), which is then used by

phosphoenolpyruvate carboxylase (PEPC) to carboxylate phosphoenolpyruvate and to generate a four-carbon molecule (10). The carbon is then released in the bundle sheath cells and used in the Calvin cycle by Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco). Unlike C_3 photosynthesizing plants, where the most highly expressed CA (βCA1) is in the chloroplast, the CA critical for the C_4 CCM is located in the cytoplasm (Figure 1.2). The same is true for plants that utilize CAM photosynthesis in order to establish a CCM (8). In C_4 and CAM plants, the CA enzymes involved in establishing their respective CCMs is highly expressed (10) as one might expect for an enzyme involved in photosynthesis, but what about C_3 plants? In C_3 plants like spinach (*Spinacia oleracea*) and mouse-ear cress (*Arabidopsis thaliana*, hereafter *Arabidopsis*), the most highly expressed CA, βCA1, is located in the stroma of the chloroplast (11) and can comprise as much as 1-2% of total leaf soluble protein (12). Given that CO_2 can freely diffuse across the chloroplast envelope to be carboxylated by Rubisco (Figure 1.2), what then is the purpose of a highly expressed CA localized to the chloroplast? In this review, the potential role(s) of chloroplast CA will be explored, highlighting the current accepted understanding of CA functionality along with the topics/areas for future research.

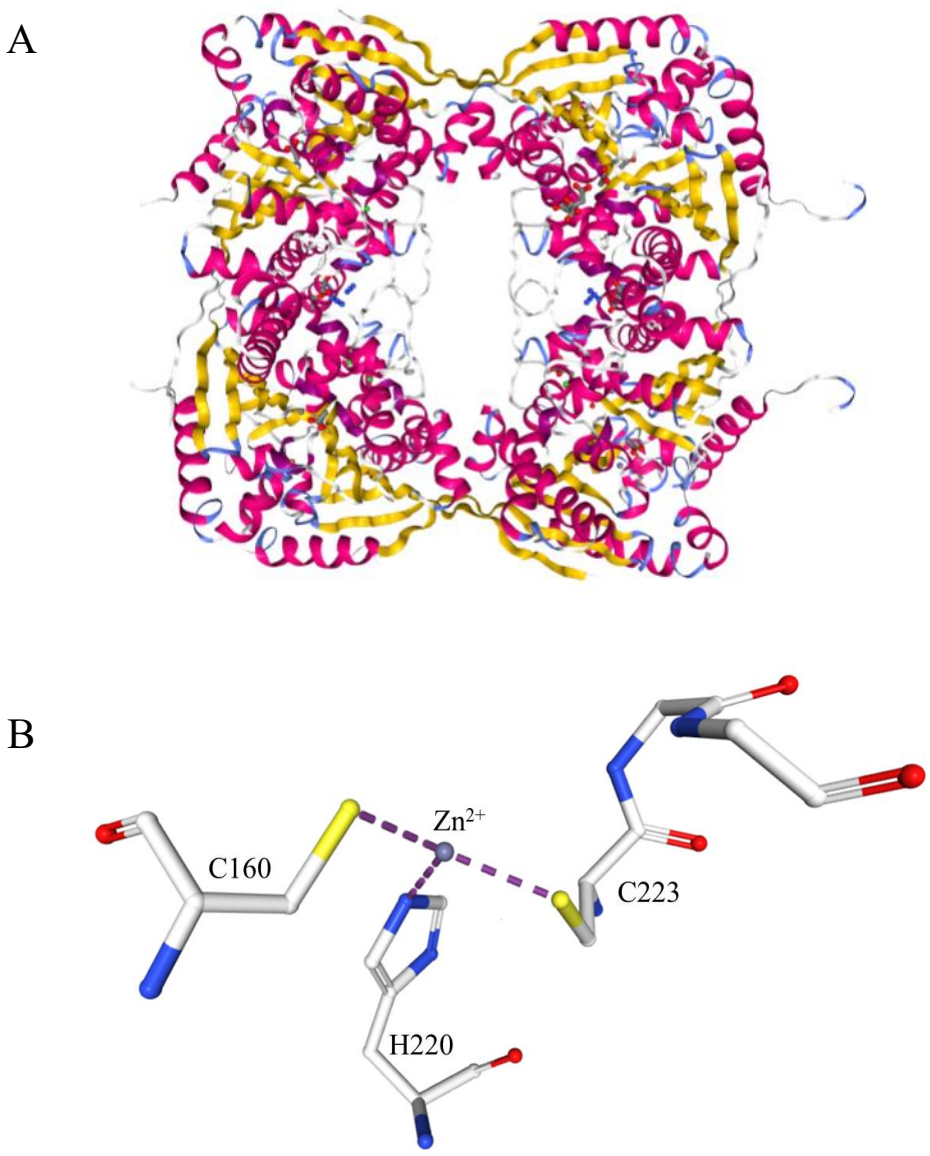

Figure 1.1 Structure of carbonic anhydrase. (A) Biological assembly X-ray crystallographic structure of a βCA1 octamer from the C_3 photosynthesizing plant *Pisum sativum* (pea), PBD file: 1EKJ (13). **(B)** Residue view of the pea βCA1 zinc ion-binding pocket with metal interactions shown in purple and relevant residues labeled. Images made with NGL Viewer.

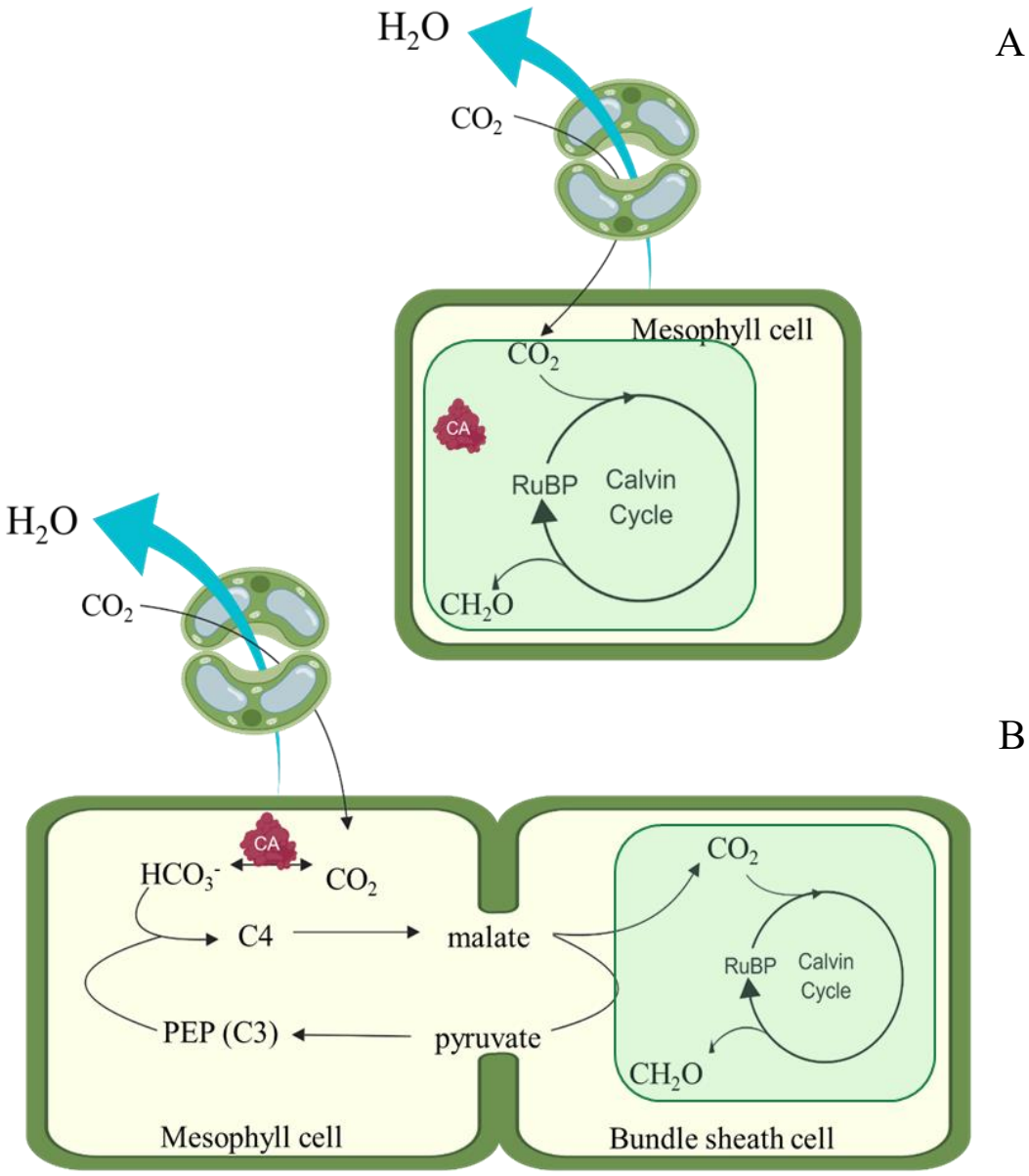

Figure 1.2 Overview of C_3 and C_4 photosynthesis. **(A)** In C_3 photosynthesis, CO_2 diffuses freely into mesophyll cells and chloroplast for use in the Calvin cycle. **(B)** C_4 photosynthesis depends on CA activity in the cytoplasm of the mesophyll cell. Image made with BioRender.

Body

Photosynthesis

Some early studies investigated the link between chloroplast-localized CA enzymes and photosynthesis. It was originally hypothesized that CA in the chloroplast of C_3 photosynthesizing plants was responsible for maintaining a CO_2 diffusion gradient across the chloroplast envelope to ensure CO_2 diffusion into the stroma and subsequent saturation of Rubisco active sites. Because the pH of the stroma (in light conditions) is about 8, the reaction in Equation 1.1 will be pushed towards the right compared to the equilibrium in the cytoplasm, which has a pH of 7.5.Therefore, at equilibrium, the CO_2 concentration in the chloroplast will be lower than in the cytoplasm, driving CO_2 diffusion into the organelle. Due to the slow, uncatalyzed interconversion rate between bicarbonate and carbon dioxide in an alkaline environment like the stroma, it was hypothesized that the bicarbonate pool would be unable to provide the CO_2 needed for Rubisco's carboxylation activity if no CA was present.

In zinc-starvation studies done with *Phaseolus vulgaris* (green bean), CA activity was reduced to below 10% of WT. This reduction had a minimal effect on growth and photosynthesis (14). The researchers also discovered that a marked decrease in zinc availability and CA activity had no effect on Rubisco levels or activation state. Of course, nutrient starvation can affect more than CA activity and the reduction in growth could be a result of any number of

unintended side effects. This study also noted that increased zinc availability had a positive effect on photosynthetic rates and chlorophyll levels in the plants' leaves.

Another early zinc-starvation study performed on *Spinacia oleracea L.* (spinach) produced similar results. Zinc-deficient leaves showed reduced photosynthetic rates (~70% of control) in high concentrations of CO_2, but no observable difference was found between zinc-starved and control plants at lower, ambient CO_2 concentrations (15). The authors concluded that equal photosynthetic rates at lower CO_2 concentrations between control plants and plants with reduced CA activity levels indicates that CA activity in the chloroplast is not necessary for CO_2 diffusion into the stroma.

Zinc-starvation studies, however, limit CA activity of all CAs regardless of cellular localization. To study the effect of CA activity reduction in the chloroplast stroma, Price *et al.* (16) used antisense RNA to target and reduce the expression of *Nicotiana tabacum* (tobacco) βCA1. They were able to generate transgenic tobacco lines with CA activity 1% of WT levels. Remarkably, only minor effects on photosynthesis and growth were observed (Figure 1.3). Antisense plants with the lowest amount of CA activity were identical to WT tobacco in terms of their gross morphology (including dry leaf weight) and showed no measurable difference in net CO_2 fixation.

Taken together, these early experiments suggested that chloroplast CA activity plays, at most, a very minor role in photosynthesis. However, none of the aforementioned studies completely eliminated CA or CA activity in the chloroplast stroma. A YFP fusion study in *A. thaliana* revealed that at least two isoforms of βCA were present in the stroma: βCA1 and βCA5 (11). When both βCA1 and βCA5 were removed from *N. tabacum*, a drastic morphological phenotype was observed at ambient CO_2 concentrations. The double mutant plants could only be rescued at extremely high CO_2 levels of 4000-9000ppm (Figure 1.3) (Chapter 2). However, PSII measurements (Φ_{PSII}), which calculate the efficiency with which absorbed protons are used to drive electron transport, showed no difference between WT and Δβ*ca1ca5* transgenic tobacco plants. There was also no evidence of any photoinhibitory damage in the double CA mutant (Chapter 2).

Past and current research of chloroplast CA has yet to discover a direct link between CA activity in the stroma and photosynthesis in C_3 plants. There is, however, increasing evidence of CA activity in the lumen of the thylakoid. Thylakoid-localized CA have been observed in the cyanobacterium *Synechocystis* sp. PCC 6803 where it is involved in CO_2 uptake (17) and in higher plants like *Zea mays* where thylakoid CA appears to be associated with PSII (18). The role CA might play in the lumen of the thylakoid, and why it is

associated with PSII, is not yet clear, but perhaps further research will clarify the relationship between luminal CA and photosynthesis.

Biotic Defense

One fascinating avenue of research on the role of stromal CAs considers their potential function in plant defense (Figure 1.3). Specifically, work done in 2002 identified βCA1 in tobacco as a salicylic acid-binding protein (SABP) (19). Salicylic acid (SA) is a plant hormone that accumulates in cells during pathogen infection and is involved in triggering both local and systematic plant defenses, which includes the hypersensitive response (HR) (20-22). Briefly, the HR is a quarantine mechanism characterized by the rapid cell death surrounding an infected area. This prevents the pathogen from spreading to neighboring cells and tissues. Researchers were able to essentially eliminate the HR of tobacco plants in which βCA1 was silenced, leading them to conclude that it plays a vital role in plant defense. Interestingly, the CA inhibitor acetazolamide did not affect βCA1's ability to bind SA, nor did SA binding affect CA activity. These data suggest that the activity of the enzyme might be separate from its role in plant defense and that the binding of SA occurs at a different site on the enzyme away from the active site.

It has also been determined that βCA1 in Arabidopsis is involved in the perception of SA and may also function as a binding partner of two proteins

required for the SA response upon infection: NON EXPRESSER OF PATHOGENESIS RELATED 1 (NPR1) and NON RECOGNITION OF BTH-4 (NRB4) (23). However, knocking out βCA1 did not reduce the plant's ability to sense SA. In fact, double, triple, and quadruple βCA mutants were required to produce a phenotypic response.

Among its multiple roles, SA is involved in closing the stomata, the obvious entry point for pathogens to infect leaf tissues (24). There is good evidence that βCA1 also plays a role in the opening and closing of the stomata. Silencing βCA1 in tobacco results in a moderate increase in stomatal conductance (a measure of CO_2 entering the leaf and water exiting), causing the stomata to remain open (16). An *Arabidopsis* double mutant (*ca1ca4*) lacking both βCA1 *and* βCA4 (found on the plasma membrane) had impaired stomatal movement and had a significant increase in stomatal density compared to WT (188 ± 13.2 stomata per mm^2 and 142.5 ± 15.4 stomata per mm^2 respectively) (25). Despite this substantial difference in stomatal conductance and density, the photosynthetic efficiency of the mutant and WT *Arabidopsis* plants was the same. Importantly however, significant stomatal phenotypes were only observed in the *ca1ca4* double mutants. Single mutants of *ca1* or *ca4* were similar to WT in both stomatal conductance and photosynthesis.

Our research found similar results to Slaymaker, *et al.* (19); transgenic tobacco lacking βCA1 showed no HR-mediated cell death when leaves were

infected with tobacco mosaic virus (TMV). The same was not true of Δ*βca5* tobacco, which still displayed an HR upon infection (Chapter 2). As predicted in previous research, Δ*βca1ca5*:35S::ΔZn-βCA1-CDS tobacco lines (a complemented double-CA-mutant expressing an inactive from of βCA1 in which its zinc-binding sites were mutated) still displayed an HR upon infection with TMV (Chapter 2). This suggests that CA activity is not required for βCA1 to play a role in the plant's defense response, indicating instead that βCA1's ability to bind to SA (and the downstream effects of that interaction) is what makes the HR possible. What, then, is the importance of the CA activity in the chloroplast and what domain/amino acid residues are critical for βCA1's SA binding capability? In Arabidopsis both βCA1 and βCA2 seem to play a role in perceiving SA levels in cells, but so far only βCA1 has been shown to be critical in SA binding and the subsequence HR, despite the high sequence identity between the proteins of the two species (Figure 1.4).

Buffering the chloroplast pH

In humans and other vertebrates, the CA activity plays an important role in the bicarbonate buffering system of the circulatory system (26, 27). Early studies suggested a similar role for chloroplast CAs: their hypothetical buffering capacity could mediate temporary changes in stromal pH resulting

from proton pumping during the light reactions of photosynthesis or the carboxylation of CO_2 by Rubisco (28, 29).

The strength of the bicarbonate buffering system in humans comes from the high concentration of bicarbonate found in the bloodstream and the speed of the CA enzyme (in this case, αCA). The bicarbonate concentration in blood (at pH 7.4) is ~24mM, which is effective at maintaining a constant pH. Currently, no consensus exists on the bicarbonate concentration in the chloroplast stroma. An early study put the estimate at ~27mM (30) in spinach chloroplasts (with pH 8), which would put the buffering capacity on par with that of humans. This same study also assigned CA activity the role of acidifying the stroma. On the other hand, a more recent review claimed the level of bicarbonate in the stroma was a mere 1mM (5). At 1mM bicarbonate, the potential buffering capacity of this system would be 1/100 less than in humans and other vertebrates.

CA capacity for regulating the pH of the chloroplast was studied in the tobacco double CA knockout mutants, Δβ*ca1ca5*, by using a chloroplast-localized pH sensitive derivative of GFP: pHluorin2. WT tobacco and CA mutants were transiently transformed with cTP-pHlourin2 (pHluorin2 with a chloroplast transit peptide) and the fluorescent intensity was measured at 405nm and 488nm. Changes in pH shift the ratio of the two measurements towards 405 (more basic) or 488 (more acidic). The double CA mutant showed an increase in 405/488 ratio compared to WT and the single mutants indicated

an increase in the stromal pH (Chapter 2). Since CO_2 can diffuse across the chloroplast envelope, this result is what one would expect if the pool of stromal bicarbonate was maintained by CA activity. Without CA, the generation of bicarbonate ions and protons would decrease significantly due to the slow uncatalyzed conversion rate, likely resulting in a more alkaline environment, which is what we observed in Δβ*ca1ca5* mutants. If the reverse hydration reaction of CA was important in providing CO_2 for Rubisco, we would predict the pH to drop as CO_2 is depleted. Ultimately, the quantitative nature of our pH data does not provide sufficient evidence for a bicarbonate buffering system. It does, however, highlight the directionality of the CA reaction in chloroplasts.

Oxidative stress tolerance

Multiple studies have examined the ability of chloroplast CAs (namely βCA1) to provide photo-oxidative stress protection. *Arabidposis* plants watered with 3 mM sodium bicarbonate had significantly reduced levels of ion leakage (a common marker for photo-oxidative stress) (31). Conversely, *βca1ca4* mutants had significantly higher levels of rapid cell death even when fertilized with sodium bicarbonate. Surprisingly, despite its high expression, there was no difference in ion leakage between WT and single-mutant *Arabidopsis βca1* plants. The WT and *βca1 Arabidopsis* also had nearly identical non-

photochemical quenching (NPQ) levels in long-day conditions, regardless of sodium bicarbonate fertilization.

When Slaymaker *et al.* (19) identified tobacco βCA1 as an SA binding protein, they also tested its antioxidant capacity in a Δ*NCE103* yeast (*Saccharomyces cerevisiae*) mutant that is sensitive to oxidative stress (32, 33). They discovered that they could complement Δ*NCE103* yeast with tobacco βCA1, allowing the yeast to grow in anaerobic conditions (in which Δ*NCE103* yeast dies). The same result was found in another paper which complemented the Δ*NCE103* yeast with *Medicago sativa* (alfalfa) chloroplast CA (32). One surprising aspect of these studies is that, although NCE103 is a CA-like protein with nearly 60% sequence similarity to tobacco βCA1 (34), it displays no CA enzymatic activity. This makes it difficult to determine if it is CA's enzymatic activity that enables it to complement Δ*NCE103* yeast. It also raises the question of whether CA activity and its antioxidant properties are independent functions.

We also investigated the antioxidant possibilities of CA in our tobacco chloroplast CA mutants using the fluorophore CellROX Green, which undergoes a conformational change upon oxidation, resulting in greater signal intensity (36). The double mutant *Δβca1ca5* showed an increase in chloroplast-localized ROS (reactive oxygen species) generation compared to WT tobacco and the two single mutants (Chapter 2). The increased ROS phenotype in the

double mutant was complemented with 35S::βCA1. An inactivated form of βCA1 with a mutated zinc-binding-site (35S::ΔZn-βCA1) still displayed high levels of ROS, suggesting that the antioxidant properties of CA in the chloroplast are dependent on CA's enzymatic activity.

Metabolism

Bicarbonate is a common source of inorganic carbon (C_i) in a large number of metabolic pathways found in plant cells (5, 36). While most organisms and even other photosynthetic organisms contain bicarbonate transporters to move bicarbonate into and out of cells/organelles, no such transporters have been discovered in higher plants; the protein family in higher plants most similar to algae bicarbonate transporters is responsible only for boric acid transport (36-38). Thus, the lack of a bicarbonate transporter in chloroplasts suggests that the bicarbonate in the stroma (and thus the bicarbonate available for bicarbonate-utilizing carboxylases) is generated by the hydration reaction of CAs. A summary of enzymes that use bicarbonate to carboxylate their substrates can be found in Table 1.1, which is adapted from DiMario *et al.* (5).

In the lipid biosynthesis pathway of higher plants, Acetyl-CoA carboxylase (ACC) requires bicarbonate as its form of C_i in order to carboxylate its substrate acetyl-CoA (39, 40). The role CA may play in providing

bicarbonate to this pathway was explored in tobacco plants treated with CA inhibitors and in antisense RNA tobacco lines in which βCA1 was silenced (41). The incorporation of acetate into total lipids was reduced by more than 50% when plants were treated with a CA inhibitor, from a lipid synthesis rate of 472.0 ± 38.4 in control tobacco to 206.4 ± 16.8 in ethoxyzolamide-treated plants. Similarly, the incorporation of acetate into the total lipids of chloroplasts decreased by 50% in the antisense tobacco lines compared to WT (Figure 1.3). Likewise, when we measured the free fatty acid (FFA) profile of WT and *Δβca1ca5* tobacco, we observed a drastic downregulation of 16:0 and 18:0 FFAs (nomenclature for FFA is *number of carbons:number of saturated bonds*) in the double mutant tobacco (Chapter 2). Both 16:0 and 18:0 are produced in the chloroplast stroma, demonstrating the role CA plays in FFA biosynthesis.

Hoang and Chapman (41) also investigated the role of CA in lipid biosynthesis in cotton (*Gossypium hirsutum*) embryos. They observed a drastic reduction in lipid synthesis in ethoxyzolamide-treated cotton plants. CA's role in embryo development may clarify other studies that found defects in reproduction. Specifically, the *Arabidopsis* βCA5 knockouts reported in Medina-Puche, *et al.* (23) were discovered to be sterile and unable to develop mature seeds. Likewise, while capable of developing seeds, the seeds produced by the *Δβca1ca5* double mutant of tobacco grown in ambient CO_2 had a greatly reduced germination rate of less than 33% (Chapter 2). When grown in high

CO_2 (thus shifting Eq. 1 to the right and providing more bicarbonate), tobacco *Δβca1ca5* seeds had a significantly improved germination rate of nearly 70%. Taken together, these studies highlight the importance of CA and the bicarbonate it generates in the lipid biosynthesis process needed for proper seedling development.

Once again, however, the amount of bicarbonate present in the chloroplasts at equilibrium has cast doubts on whether it is necessary for these pathways. Conservative estimates of bicarbonate concentration in the chloroplast stroma are well below the amount needed to saturate enzymes like Acetyl-CoA Carboxylase (ACC) (5). While the lack of CA would decrease the bicarbonate supply for these reactions, many are already less than 10% saturated even when CA is present. This might be one of the many reasons why chloroplast CA activity can be reduced to 1% of wild-type with nearly no harmful effects on plant development (16). Predictions that put the bicarbonate concentration at the lower end would point to the importance of CA-derived bicarbonate in pathways with enzymes that have a high K_m. In Table 1, the enzyme 5-aminoimidazole ribonucleotide carboxylase would be the most affected by the a reduction in bicarbonate availability due to its K_m of 23.0mM (42)., Nevertheless, disruptions in the purine synthesis pathway related to CA activity were found in *Arabidopsis* β*ca2ca4* mutants in which both βCA2 and βCA4 are localized to the cytoplasm (43).

If removing CA from the chloroplast stroma results in bicarbonate becoming a limiting resource in multiple biosynthetic pathways, we would expect to see an increase in the stromal pH as the available bicarbonate is used up and is only slowly replaced by the natural hydration of CO_2. This is what we observed in pH measurements of Δβ*ca1ca5* tobacco mutants (Chapter 2).

Pathway	Enzyme	Enzyme location	Km (mM)
Photosynthesis (C_3)	Rubisco	Chloroplast	0.009-0.120
Photosynthesis (C_4)	Phosphoenolpyruvate Carboxylase	Cytoplasm	0.027-0.180
Fatty Acid Synthesis	Acetyl-CoA Carboxylase	Chloroplast	0.9-2.5
Purine Synthesis	5-aminoimidazole Ribonucleotide Carboxylase	Chloroplast	23

Table 1.1 Select plant enzymes that use C_i to carboxylate their substrates.

Adapted and simplified from DiMario et al. (5) with Km values given in the mM range from the following references: (40, 42, 44-46).

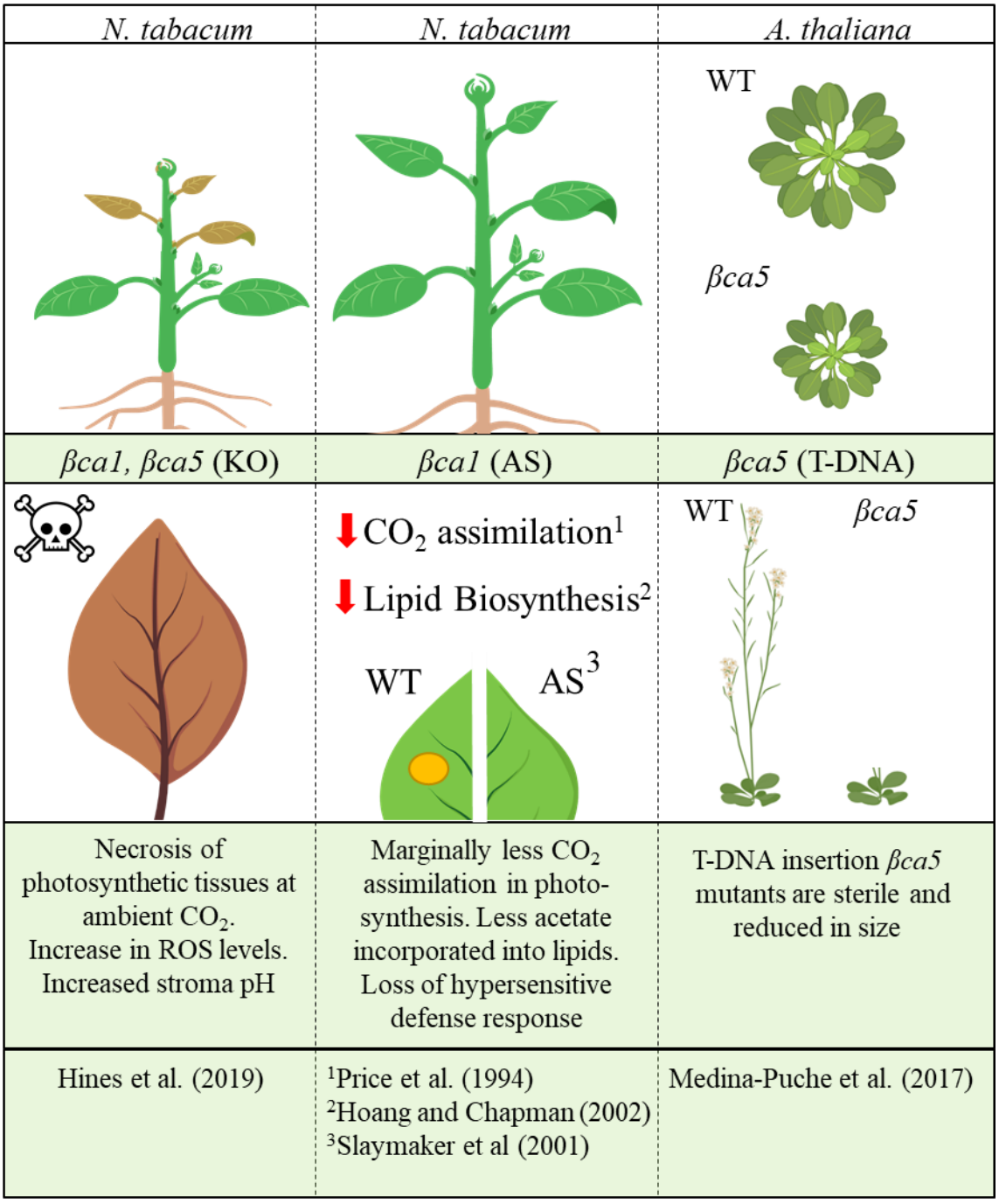

Figure 1.3 Summary of select chloroplast CA mutants. Summary of CA knockouts (KO), antisense RNA (AS), and T-DNA insertion (T-DNA) mutants in tobacco and *Arabidopsis*, with descriptions of major phenotypes. Images made in BioRender.

Discussion and Perspectives

Despite decades of research, there is still no straightforward understanding of the exact functions of chloroplast CA. The enzymes remain a sort of "jack of all (plant) trades." Yet, this explanation does not explain many of the observations researchers have made in previous studies, like the fact that although CA can comprise upwards of 2-5% of all soluble protein in the leaf, the gross morphologies of the plants and their photosynthetic performances remain unaffected when the amount of CA is considerably reduced (9, 12, 16). If only a small amount of chloroplast CA is needed, why do so many C_3 plants spend so much nitrogen and energy expressing it in such large quantities?

Furthermore, studies are complicated by the diversity of CAs even among C_3 plants. For example, tobacco contains two CAs in its chloroplasts, βCA1 and βCA5 (Chapter 2)(19), while *Arabidopsis* has three chloroplast CAs: βCA1, βCA5 and αCA1, the latter of which takes the unique route to the chloroplast by first following the secretory pathway (11, 47), despite the similarities between the CAs of tobacco and Arabidopsis in which the αCA1 amino acid sequences are nearly 50% identical to each other (Table 1.2).

Homology does not always translate to shared functions, especially with higher plant CAs. While βCA1 and βCA2 in Arabidopsis are fairly similar with a shared sequence identity of 71.6% (Figure 1.4), they appear to have separate functions and only βCA1 appears to play a substantial role in plant defense via

the HR (19). In a phylogenetic tree, Arabidopsis βCA1 and βCA2 even group together (Figure 1.5). βCA1 and βCA5 are both found in the chloroplast and, as the single Δ*βca1* and Δ*βca5* mutant tobacco lines show, their functions are largely redundant in terms of development, but they are only 34.55% identical (Figure 1.4).

Confusingly, even CA homologs of the same name and same localization can have different functions. When βCA5 is knocked out in *Arabidopsis*, the resulting mutants are slow-growing and sterile, with T_0 plants being unable to produce T_1 seeds (23). Meanwhile, tobacco Δ*βca5* mutants grow and produce viable seeds in the same manner as WT tobacco (Chapter 2). The phenotype observed in *βca5 Arabidopsis* is especially unexpected given that the CA activity levels between the mutant and WT *Arabidopsis* are essentially equal (23)

Arabidopsis CA	Identifier	Location	Tobacco CA		Identifier	Location	%ID
βCA1	AT3G01500	Chloroplast	βCA1	Ns	NP_001313031	Chloroplast	69.74
			βCA1	Nt	XP_016504732	Chloroplast	60.95
βCA5	AT4G33580	Chloroplast	βCA5	Ns	XP_016482109	Chloroplast	49.17
			βCA5	Nt	XP_016446541	Chloroplast	47.85
αCA1	AT3G52720	Chloroplast	αCA1		XP_016483347	PM	47.81

Table 1.2 Sequence similarities of Arabidopsis and tobacco CAs.
Arabidopsis protein identifiers come from The Arabidopsis Information Resource (TAIR) and tobacco identifies come from the National Center for Biotechnology Information (NCBI) accession numbers.

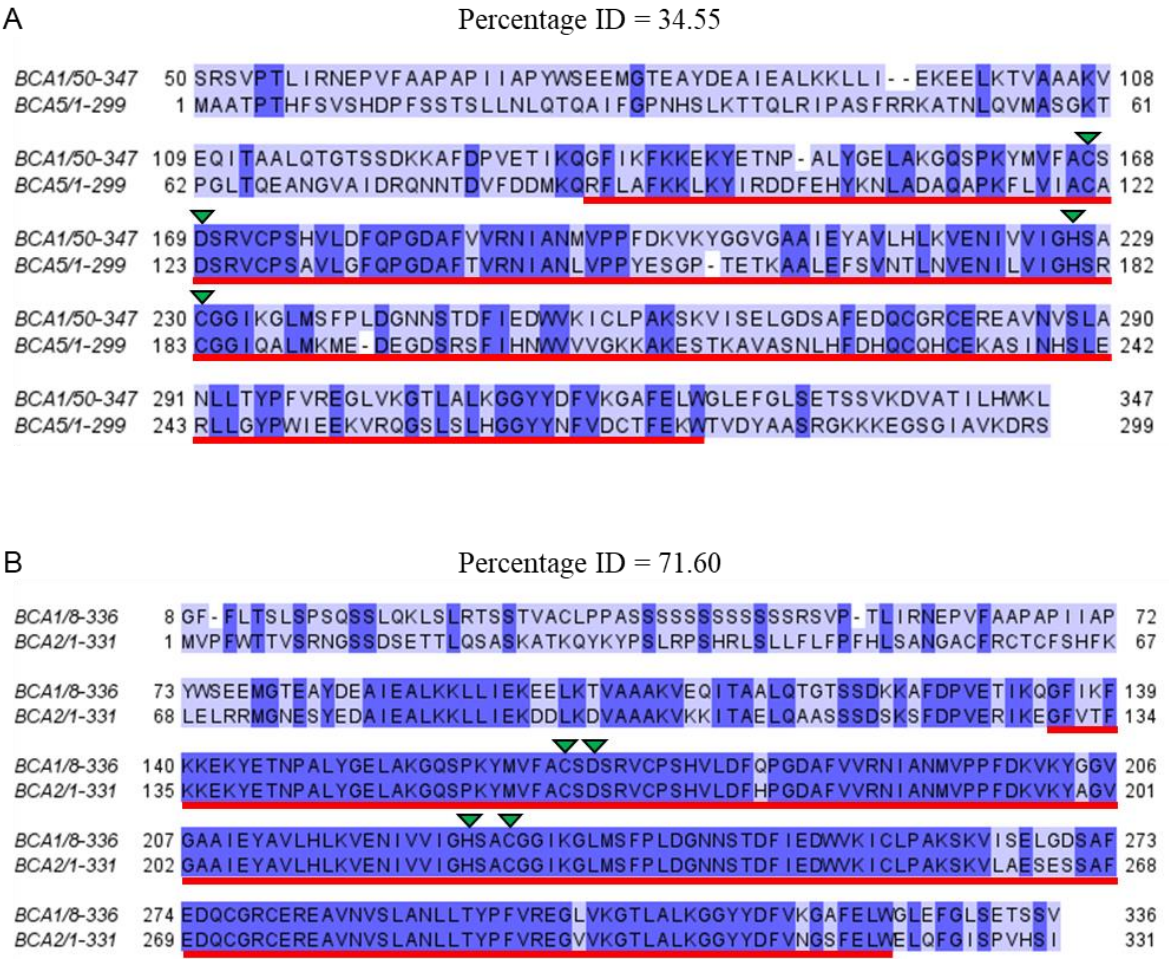

Figure 1.4. Protein sequence alignment between Arabidopsis beta CAs. (A) Pairwise alignment between Arabidopsis βCA1 (AT3G01500) and βCA5 (AT4G33580) amino acid sequences shows a 34.55% identity. **(B)** Pairwise alignment between Arabidopsis βCA1 (AT3G01500) and βCA2 (AT5G14740) amino acid sequences shows a 71.6% identity. Dark blue = Identical residues. Red line = Conserved beta CA superfamily (Accession: cl00391). Green triangles = Zinc binding residues.

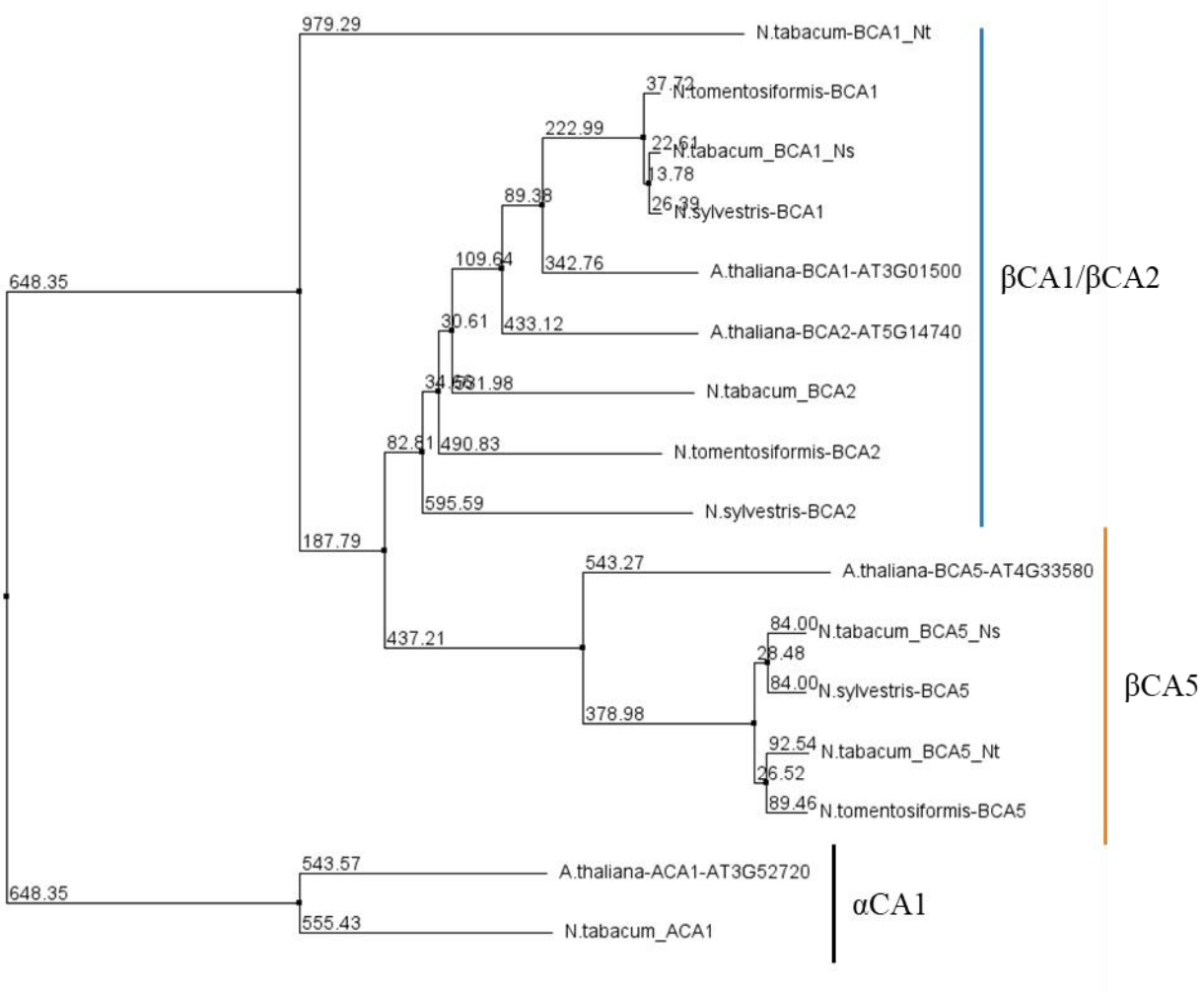

Figure 1.5. Phylogeny of CAs from Arabidopsis and Nicotiana. Neighbor joining calculations using BLOSUM62 in the Jalview software of CAs found in the chloroplast (βCA1, βCA5, and αCA1) and the highly similar βCA2 from Arabidopsis and three species of *Nicotiana* sp. (Ns) and (Nt) indicate homology of *N. tabacum* protein to either *N. sylvestris* or *N. tomentosiformis*, respectively.

The divergent, unrelated roles a single CA enzyme can play likewise make it difficult to assign concrete function to CAs in general, and chloroplast CAs, specifically. CA is obviously known and characterized by the reversible hydration reaction it rapidly catalyzes, but βCA1, for example, is a SA binding protein that is necessary to trigger the hypersensitive response crucial to plant defense. In *Arabidopsis*, βCA1 and βCA5 are involved in the perception of SA levels in leaves (23), although mutant *βca5* tobacco plants still exhibit an HR phenotype when infected with TMV (Chapter 2).

While elucidating the vital role(s) of chloroplast CAs is challenging, the research is critical for scientists aiming to engineer chloroplasts to improve photosynthesis. It is important to know where CAs are localized, what drives their expression, and what their distinct responsibilities are before researchers can most effectively engineer C_4 photosynthesis into C_3 plants in order to improve crop yields (48, 49). As we have seen in Δβ*ca1ca5* tobacco mutants expressing a cytoplasmically-localized βCA1 (Δ62-βCA1), CA activity in the cytoplasm cannot complement a plant lacking stromal CA.

Likewise, there have also been considerable research effort in the ongoing attempts to engineer carboxysomes from cyanobacteria (which contributes to its CCM) into the chloroplasts of C_3 crop plants in order to improve photosynthesis and nitrogen use efficiency (50). The Rubisco from *Synechococcus elongatus* is significantly faster than native plant Rubisco (51)

and the carboxysome shell it is encapsulated within concentrates CO_2 around its active sites, while excluding O_2 and oxygenation. One of the challenges for the carboxysome engineering project is the presence of CA (and CA activity) in the chloroplast stroma. When, for example, human αCA1 (HCA1) was expressed in the cytoplasm of *S. elongatus* (which would generate CO_2 *outside* of the carboxysome shell, disrupting the CCM), the resulting mutants were slow growing and required elevated CO_2 levels in order to survive (52). The human HCA1 did not interfere with the assembly of the carboxysome, but rather with the CCM. If carboxysomes are to be successfully engineered into chloroplasts and function as part of a larger, synthetic CCM, native CAs in the stroma will need to be removed. But as demonstrated above, completely removing CA from the stroma results in a drastic phenotype. Nevertheless deactivated βCA1 can still contribute to plant defense and the bicarbonate needed for the multiple pathways listed in Table 1.1 could potentially be provided by a bicarbonate pump engineered into the chloroplast envelope (Chapter 2)(50).

In sum, the study of plant CAs has elucidated various biological functions and provided insights in the evolutionary history of C_3 and C_4 photosynthesis (53). Studying this class of enzyme will also contribute to the ongoing efforts to modify photosynthesis and improve crop yields for an ever-growing world population

References

1. D. S. Goodsell, Carbonic Anhydrase. *RCSB Protein Data Bank* 10.2210/rcsb_pdb/mom_2004_1 (2004).
2. Claudiu T. Supuran, Structure and function of carbonic anhydrases. *Biochemical Journal* **473**, 2023-2032 (2016).
3. S. Lindskog, Structure and mechanism of carbonic anhydrase. *Pharmacology & Therapeutics* **74**, 1-20 (1997).
4. W. B. Frommer, CO_2mmon Sense. *Science* **327**, 275-276 (2010).
5. R. J. DiMario, M. C. Machingura, G. L. Waldrop, J. V. Moroney, The many types of carbonic anhydrases in photosynthetic organisms. *Plant Science* **268**, 11-17 (2018).
6. H. Park, P. J. McGinn, F. M. M. Morel, Expression of cadmium carbonic anhydrase of diatoms in seawater. *Aquatic Microbial Ecology* **51**, 183-193 (2008).
7. X. Ren, S. Lindskog, Buffer dependence of CO2 hydration catalyzed by human carbonic anhydrase I. *Biochim Biophys Acta* **1120**, 81-86 (1992).
8. R. J. DiMario, H. Clayton, A. Mukherjee, M. Ludwig, J. V. Moroney, Plant carbonic anhydrases: Structures, locations, evolution, and physiological roles. *Molecular Plant* **10**, 30-46 (2017).
9. M. R. Badger, G. D. Price, The role of carbonic anhydrase in photosynthesis. *Annual Review of Plant Physiology and Plant Molecular Biology* **45**, 369-392 (1994).
10. S. Von Caemmerer *et al.*, Carbonic anhydrase and C4 photosynthesis: a transgenic analysis. *Plant, Cell and Environment* **27**, 697-703 (2004).
11. N. Fabre, I. M. Reiter, N. Becuwe-Linka, B. Genty, D. Rumeau, Characterization and expression analysis of genes encoding alpha and beta carbonic anhydrases in Arabidopsis. *Plant Cell Environ* **30**, 617-629 (2007).
12. K. Okabe, S.-Y. Yang, M. Tsuzuki, S. Miyachi, Carbonic anhydrase: Its content in spinach leaves and its taxonomic diversity studied with anti-spinach leaf carbonic anhydrase antibody. *Plant Science Letters* **33**, 145-153 (1984).
13. M. S. Kimber, The active site architecture of *Pisum sativum* beta -carbonic anhydrase is a mirror image of that of alpha -carbonic anhydrases. *The EMBO Journal* **19**, 1407-1418 (2000).
14. G. E. Edwards, A. K. Mohamed, Reduction in carbonic anhydrase activity in zinc deficient leaves of *Phaseolus vulgaris* L.1. *Crop Science* **13** (1973).
15. P. J. Randall, D. Bouma, Zinc Deficiency, Carbonic anhydrase, and photosynthesis in leaves of spinach. *Plant Physiology* **52**, 229-232 (1973).

16. G. D. Price *et al.*, Specific reduction of chloroplast carbonic anhydrase activity by antisense RNA in transgenic tobacco plants has a minor effect on photosynthetic CO_2 assimilation. *Planta* **193**, 331-340 (1994).

17. N. Sun *et al.*, A thylakoid-located carbonic anhydrase regulates CO_2 uptake in the cyanobacterium *Synechocystis* sp. PCC 6803. *New Phytologist* **222**, 206-217 (2018).

18. Y.-K. Lu, A. J. Stemler, Extrinsic photosystem ii carbonic anhydrase in maize mesophyll chloroplasts. *Plant Physiology* **128**, 643-649 (2002).

19. D. H. Slaymaker *et al.*, The tobacco salicylic acid-binding protein 3 (SABP3) is the chloroplast carbonic anhydrase, which exhibits antioxidant activity and plays a role in the hypersensitive defense response. *Proceedings of the National Academy of Sciences* **99**, 11640-11645 (2002).

20. K. E. Hammond-Kosack, J. D. Jones, Resistance gene-dependent plant defense responses. *The Plant Cell* **8**, 1773-1791 (1996).

21. J. Durner, J. Shah, D. F. Klessig, Salicylic acid and disease resistance in plants. *Trends in Plant Science* **2**, 266-274 (1997).

22. D. M. A. Dempsey, J. Shah, D. F. Klessig, Salicylic acid and disease resistance in plants. *Critical Reviews in Plant Sciences* **18**, 547-575 (2010).

23. L. Medina-Puche *et al.*, beta-carbonic anhydrases play a role in salicylic acid perception in Arabidopsis. *PLoS One* **12**, e0181820 (2017).

24. M. A. R. Khokon *et al.*, Involvement of extracellular oxidative burst in salicylic acid-induced stomatal closure in Arabidopsis. *Plant, Cell & Environment* **34**, 434-443 (2011).

25. H. Hu *et al.*, Carbonic anhydrases are upstream regulators of CO_2-controlled stomatal movements in guard cells. *Nature Cell Biology* **12**, 87-93 (2009).

26. N. U. Meldrum, F. J. W. Roughton, Carbonic anhydrase. Its preparation and properties. *The Journal of Physiology* **80**, 113-142 (1933).

27. B. J. Krieg, S. M. Taghavi, G. L. Amidon, G. E. Amidon, *In vivo* predictive dissolution: Transport analysis of the co_2 , bicarbonate *in vivo* buffer system. *Journal of Pharmaceutical Sciences* **103**, 3473-3490 (2014).

28. B. S. Jacobson, F. Fong, R. L. Heath, Carbonic anhydrase of spinach. *Plant Physiology* **55**, 468-474 (1975).

29. H. Pfanz, U. Heber, Buffer capacities of leaves, leaf cells, and leaf cell organelles in relation to fluxes of potentially acidic gases. *Plant Physiology* **81**, 597-602 (1986).

30. U. Wagner, J. Kolbowski, V. Oja, A. Laisk, U. Heber, pH homeostasis of the chloroplast stroma can protect photosynthesis of leaves during the influx of potentially acidic gases. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1016**, 115-120 (1990).

31. J. Dąbrowska-Bronk *et al.*, β-carbonic anhydrases and carbonic ions uptake positively influence Arabidopsis photosynthesis, oxidative stress tolerance and growth in light dependent manner. *Journal of Plant Physiology* **203**, 44-54 (2016).

32. R. Götz, A. Gnann, F. K. Zimmermann, Deletion of the carbonic anhydrase-like geneNCE103 of the yeast *Saccharomyces cerevisiae* causes an oxygen-sensitive growth defect. *Yeast* **15**, 855-864 (1999).

33. A. E. Cleves, A new pathway for protein export in *Saccharomyces cerevisiae*. *The Journal of Cell Biology* **133**, 1017-1026 (1996).

34. R. Lehneck, S. Pöggeler, A matter of structure: structural comparison of fungal carbonic anhydrases. *Applied Microbiology and Biotechnology* **98**, 8433-8441 (2014).

35. E. Lozoya-Gloria, I. Cornejo-Corona, H. Thapa, D. Browne, T. Devarenne, ROS detection in *Botryococcus braunii* colonies with CellROX Green reagent. *Bio-Protocol* **7** (2017).

36. C. Poschenrieder *et al.*, Transport and use of bicarbonate in plants: Current knowledge and challenges ahead. *International Journal of Molecular Sciences* **19** (2018).

37. J. Takano, K. Miwa, T. Fujiwara, Boron transport mechanisms: Collaboration of channels and transporters. *Trends in Plant Science* **13**, 451-457 (2008).

38. J. Takano *et al.*, Polar localization and degradation of Arabidopsis boron transporters through distinct trafficking pathways. *Proceedings of the National Academy of Sciences* **107**, 5220-5225 (2010).

39. C. Alban, P. Baldet, R. Douce, Localization and characterization of two structurally different forms of acetyl-CoA carboxylase in young pea leaves, of which one is sensitive to aryloxyphenoxypropionate herbicides. *Biochemical Journal* **300**, 557-565 (1994).

40. B. J. Nikolau, J. C. Hawke, Purification and characterization of maize leaf acetyl-coenzyme a carboxylase. *Archives of Biochemistry and Biophysics* **228**, 86-96 (1984).

41. C. V. Hoang, K. D. Chapman, Biochemical and Molecular Inhibition of Plastidial Carbonic anhydrase reduces the incorporation of acetate into lipids in cotton embryos and tobacco cell suspensions and leaves. *Plant Physiology* **128**, 1417-1427 (2002).

42. S. M. Firestine, V. J. Davisson, Carboxylases in de novo purine biosynthesis. characterization of the *Gallus gallus* bifunctional enzyme. *Biochemistry* **33**, 11917-11926 (2002).

43. R. J. DiMario *et al.*, The cytoplasmic carbonic anhydrases βCA2 and βCA4 are required for optimal plant growth at low CO_2. *Plant Physiology* **171**, 280-293 (2016).

44. D. B. Jordan, W. L. Ogren, Species variation in the specificity of ribulose biphosphate carboxylase/oxygenase. *Nature* **291**, 513-515 (1981).

45. P. H. Reibach, C. R. Benedict, Fractionation of stable carbon isotopes by phosphoenolpyruvate carboxylase from C4 plants. *Plant Physiology* **59**, 564-568 (1977).

46. B. J. Nikolau, J. B. Ohlrogge, E. S. Wurtele, Plant biotin-containing carboxylases. *Archives of Biochemistry and Biophysics* **414**, 211-222 (2003).

47. A. Villarejo *et al.*, Evidence for a protein transported through the secretory pathway en route to the higher plant chloroplast. *Nature Cell Biology* **7**, 1224-1231 (2005).

48. M. L. Schuler, O. Mantegazza, A. P. M. Weber, Engineering C4 photosynthesis into C3chassis in the synthetic biology age. *The Plant Journal* **87**, 51-65 (2016).

49. L. Wang *et al.*, Comparative analyses of C4 and C3 photosynthesis in developing leaves of maize and rice. *Nature Biotechnology* **32**, 1158-1165 (2014).

50. G. D. Price, S. M. Howitt, Towards turbocharged photosynthesis. *Nature* **513**, 497-498 (2014).

51. M. T. Lin, A. Occhialini, P. J. Andralojc, M. A. J. Parry, M. R. Hanson, A faster Rubisco with potential to increase photosynthesis in crops. *Nature* **513**, 547-550 (2014).

52. G. D. Price, M. R. Badger, Expression of human carbonic anhydrase in the cyanobacterium *Synechococcus* PCC7942 creates a high CO_2-requiring phenotype. *Plant Physiology* **91**, 505-513 (1989).

53. M. Ludwig, Evolution of carbonic anhydrase in C4 plants. *Current Opinion in Plant Biology* **31**, 16-22 (2016).

Chapter 2

Chloroplast carbonic anhydrase activity in C_3 photosynthesizing plants is essential for normal development but not for photosynthesis[5]

ABSTRACT

The enzyme carbonic anhydrase (CA) catalyzes the interconversion of bicarbonate (HCO_3^-) with carbon dioxide (CO_2) and water. We identified two stromal CAs, βCA1 and βCA5, and produced CRISPR/Cas9 mutants affecting both genes. While the single knockout lines *Δβca1* and *Δβca5* had no striking phenotypic differences compared to WT plants, *Δβca1ca5* leaves developed abnormally and exhibited large necrotic lesions, even when supplied with sucrose. Leaf development of *Δβca1ca5* plants normalized at the high CO_2 concentration of 9000ppm. High CO_2-grown *Δβca1ca5* mutants had no measurable defect in photosystem II efficiency when measured at ambient CO_2 even though emerging *Δβca1ca5* leaves show an upsurge in chloroplast reactive oxygen species (ROS). *Δβca1ca5* seedling germination and development is negatively affected when seedling development occurs at ambient CO_2. A series of complementation experiments using altered forms of βCA1 were carried out

[5] Chapter presented as manuscript in preparation: Hines, KM., Edgeworth, KN., Owens, TG., and Hanson, MR. (2019) *Chloroplast carbonic anhydrase activity in C3 photosynthesizing plants is essential for normal development but not for photosynthesis*

in Cas9-lacking (cas9$^-$) *Δβca1ca5* plants. Constructs expressing full length βCA1 and βCA5 proteins complemented the *Δβca1ca5* mutation, but inactivated (ΔZn-βCA1) and cytoplasm-localized (Δ62-βCA1) forms of βCA1 failed to reverse the mutant phenotype. When infected with tobacco mosaic virus (TMV) *Δβca1* and *Δβca1ca5* tobacco failed to show the hypersensitive response (HR), while expression of ΔZn-βCA1 restored the response. Thus, stromal CAs play major roles in plant development and defense.

Introduction

The zinc-containing metalloenzyme carbonic anhydrase (CA, EC 4.2.1.1) catalyzes the reversible hydration of carbon dioxide with bicarbonate and protons (Equation 1.1). CA greatly increases the rate by which carbon dioxide and bicarbonate reach equilibrium by several orders of magnitude. CAs are a ubiquitous and diverse group of enzymes that are found in all living organisms. Seven distinct families of CAs exist , α-, β-, γ-, δ-, ζ-, η-, and ɵ-CAs, all of which vary greatly in their sequences and structures but carry out same reversible hydration reaction in a remarkable example of convergent evolution (1).AS a result, different organisms will have specific families of CAs to carry out various

functions. In mammals, for example, αCAs are responsible for the vital bicarbonate buffering system in blood (2).

Vascular land plants contain only α-, β-, and γ- CAs (3), which are responsible for a myriad of physiological roles depending on the plant and the type of photosynthesis it carries out. In C_4-photosynthesizing plants, highly expressed βCAs in the cytoplasm play a critical role in the C_4 carbon concentration mechanism (CCM) by converting CO_2 to HCO_3^-, which then can be fixed by Phosphoenolpyruvate Carboxylase (PEPCase) (4). The importance of CAs for proper photosynthetic function in C_4 plants is therefore well understood (5).

In C_3 plants, CAs might potentially be involved in photosynthesis by quickly establishing the equilibrium concentration of CO_2 and bicarbonate as CO_2 is consumed by carbon fixation and protons are released. In light in wild-type chloroplasts, the stroma becomes more alkaline, approaching pH8 from an initial pH of approximately 7 (6). The higher pH will drive the reaction to the right, favoring formation of bicarbonate. Theoretically CA could be needed to reestablish the equilibrium, because the spontaneous reaction is very slow. A chloroplast lacking CA would be expected to exhibit lower pH than a wild-type chloroplast *if* CA activity is required to produce CO_2 from bicarbonate for carbon fixation.

Research on CAs in C_3 plants has resulted in the assignment of several roles to the enzymes. The chloroplasts of Arabidopsis are known to contain at least two stromal CAs: βCA1 and βCA5 (7). In C_3 photosynthesizing plants, the chloroplast-localized βCA1 is the most highly expressed CA in leaf tissue, with nearly 50 times the amount of RNAseq reads and 13 times the amount of EST counts compared to other plant CAs (8). Tobacco βCA1 is a salicylic acid (SA) binding protein and participates in plant defense. Silencing of βCA1 in tobacco resulted in a reduction in the hypersensitive response during a *Pto-avrPto* interaction (9) and, more recently, multiple βCAs have been shown to play a role in the perception of SA in *Arabidopsis thaliana* (hereafter Arabidopsis) (10). Mutating *both* the chloroplast localized βCA1 and plasma membrane localized βCA4 in Arabidopsis caused a reduction in stomatal CO_2 response (11, 12). In another study of the *βca1ca4* Arabidopsis mutant, researchers observed an intolerance to oxidative stress, suggesting that CAs play a role in cell death homeostasis during light stress (13). Furthermore, CAs may also play an important function in certain cellular biosynthesis pathways, as demonstrated in Arabidopsis *βca2ca4* mutants, which showed a significant decrease in aspartate that was rescued when plants were grown in elevated CO_2 levels (8). Transgenic tobacco plants carrying an antisense construct reduced CA

expression to 5% of the wild-type level, resulting in normal plant development despite altered lipid biosynthesis (14).

Despite the fact that βCAs in C_3 plants can constitute as much as 1-2% of total leaf protein (15), C_3 plants appear to have little need for βCAs for photosynthesis (5). Indeed, in antisense transgenic lines, where chloroplast CA activity was reduced to 1% of WT levels in *Nicotiana tabacum* (tobacco), there was no significant reduction in the CO_2 assimilation rate (16). However, βCAs have rapid catalytic activity, so that even as little as 1% CA could potentially be adequate for many of its functions in the stroma. If C_3 plants are able to perform photosynthesis normally with only 1% of WT CA activity, why do they spend so much nitrogen and energy expressing so much? The idea that C_3 plants are already making more than enough CA is also supported by studies that found no improve in photosynthesis (or any changes in other physiological functions) when CA was overexpressed in tobacco (17).

Interest in the role of stromal CAs has recently been reawakened because of efforts to install carbon-concentrating mechanisms into C_3 plant chloroplasts (18, 19). In order to supply a high CO_2 environment to Rubisco within carboxysomes of cyanobacteria, CA is absent from the cytoplasm, instead confined to the microcompartment (20, 21). Incorporating a operational carboxysome within

chloroplasts will require removal of all CA in the stroma (22), which is the functional equivalent of the cytoplasm of cyanobacteria. However, the consequences of a complete lack of stromal CA activity has not previously been investigated.

Here we identify βCA1 and βCA5 as the tobacco chloroplast stromal CAs and use the CRIPSR/Cas9 system to generate knockout lines, thus removing all CA activity from the stroma. We discovered that the absence of stromal CA activity had unexpected detrimental effects on leaf, floral bud and seed development and increased the reactive oxygen species and increased, rather than decreased, the pH within chloroplasts. We demonstrate that the developmental phenotypes can be reversed by growing the mutant plants in high (9000ppm) CO_2, which is known to result in increased non-enzymatic production of bicarbonate.

RESULTS

Tobacco CA enzymes βCA1 and βCA5 localize to the chloroplast stroma

The coding sequences (CDS) of 3 candidate tobacco CA genes were selected based on TargetP sequence-based predictions (βCA1, and βCA5) and previous localization studies of CA homologs in Arabidopsis (αCA1). A series of YFP fusion experiments were carried out to determine which of these tobacco CA enzymes localize to the chloroplast stroma. In Arabidopsis, βCA1 and βCA5

have been observed to localize to the chloroplast stroma (7). βCA1 in Arabidopsis and tobacco are 76% identical, and the two βCA5 genes are 59% identical. Tobacco YFP-fused βCA1 and βCA5 were both found to be localized to the chloroplast stroma (Figure 2.1A). Interestingly, βCA1-YFP was sometimes found in the cytoplasm or in a "ring" around the chloroplast when imaged 48 hours after inoculation (Supplemental Figure S2.1A). This pattern is consistent with a bottleneck in the chloroplast import machinery due to a result of the transient transformation. There are 4 predicted isoforms for βCA5 in tobacco (Accession numbers XP_016482110, XP_016482109, XP_016446542, and XP_016446541), two for each of the two tobacco progenitor *N. sylvestris* and *N. tomentosiformis* alleles, but reverse transcription of tobacco leaf RNA yielded only 1 isoform (see methods), which was used in our experiments.

The tobacco αCA1 is predicted to be a secretory pathway protein with a transmembrane sequence near the N-terminus, but it is reported to be localized to the chloroplast stroma in Arabidopsis (23). YFP-fused tobacco αCA1 did not localize to the chloroplast stroma and instead appeared to be on the plasma membrane, which is consistent with the software's prediction (Supplemental figures S2.1A and S2.1B). Taken together, YFP localization indicates that βCA1 and βCA5 are the CA enzymes in the stroma of tobacco chloroplasts.

Targeting tobacco βca1 and βca5 with CRISPR/Cas9

The CRISPR/Cas9 system was used to generate mutations in the tobacco *βca1* and *βca5* genes to create three transgenic lines: Δ*βca1*, Δ*βca5,* and the double knockout Δ*βca1ca5*. Tobacco is an allotetraploid (a hybrid of *Nicotiana sylvestris and Nicotiana tomentosiformis*), so there are two homologs for each of the CA genes. The sgRNAs were designed to target the sequences of both forms of the CA homologs. Two sgRNAs (less than <100bp apart) were used to target each CA gene to create large deletions between the two sgRNA target loci that would be visible by PCR of the gDNA in the T_0 generation of plants (Supplemental Figure S2.2).

The Δβ*ca1* transgenic line contains a 2bp deletion in the mRNA sequence of one homolog of the βca1 gene and a 52bp deletion in the other (Supplemental Figure S2.3). Both deletions created a frame-shift mutation, resulting in an early stop codon in the coding sequences. Δ*βca5* mutants contain a 2bp and 31bp deletion in the mRNA sequence of the two βca5 homologs, both of which result in frame-shift mutations. The βca1 homologs in the double mutant Δ*βca1ca5* comprise of a 1bp insertion or a 2bp deletion and the βca5 homologs both have large, 31bp deletion between the two sgRNA target sites (Supplemental Figure S2.3). The sequencing data demonstrates the creation of frame-shift mutations in the targeted CA genes in all the CRISPR/Cas9-generated transgenic lines.

The gross morphology of Δβ*ca1* and Δβ*ca5* is similar to WT tobacco when the plants are grown at ambient CO_2 concentrations (Figure 2.1B). In contrast, Δ*βca1ca5* displays a dramatic developmental phenotype, which results in pale and shriveled leaves. The plants in Figure 2.1B were first grown at 9000ppm CO_2 for 4 weeks before being transferred to ambient CO_2 for two weeks and then imaged. The white arrows highlight Δ*βca1ca5* leaves which developed in high CO_2 before being transferred.

Carbonic anhydrase activity in mutant lines

The CA activity of the lines were measured in whole leaf homogenate from plants growing at ambient CO_2 in 16hr/day light on leaves which expanded under high CO_2 conditions. We used the CA activity assay described in Medina-Puche, *et al.* (10) with modifications (see methods). This assay measures the activity of CAs in all subcellular locations of the tissue used, not only activity within the chloroplasts. Δ*βca1* and Δ*βca1ca5*, the two lines in which βCA1 was mutated, show greatly decreased levels of CA activity (Figure 2.1C), while Δ*βca5* CA activity showed no reduction. As expected, most of the CA activity found in leaves appears to come from the highly expressed βCA1 enzyme, with the Δ*βca1* transgenic line showing an 87% reduction in CA activity (0.055 Units) compared to WT (0.394 Units) (Figure 2.1C).

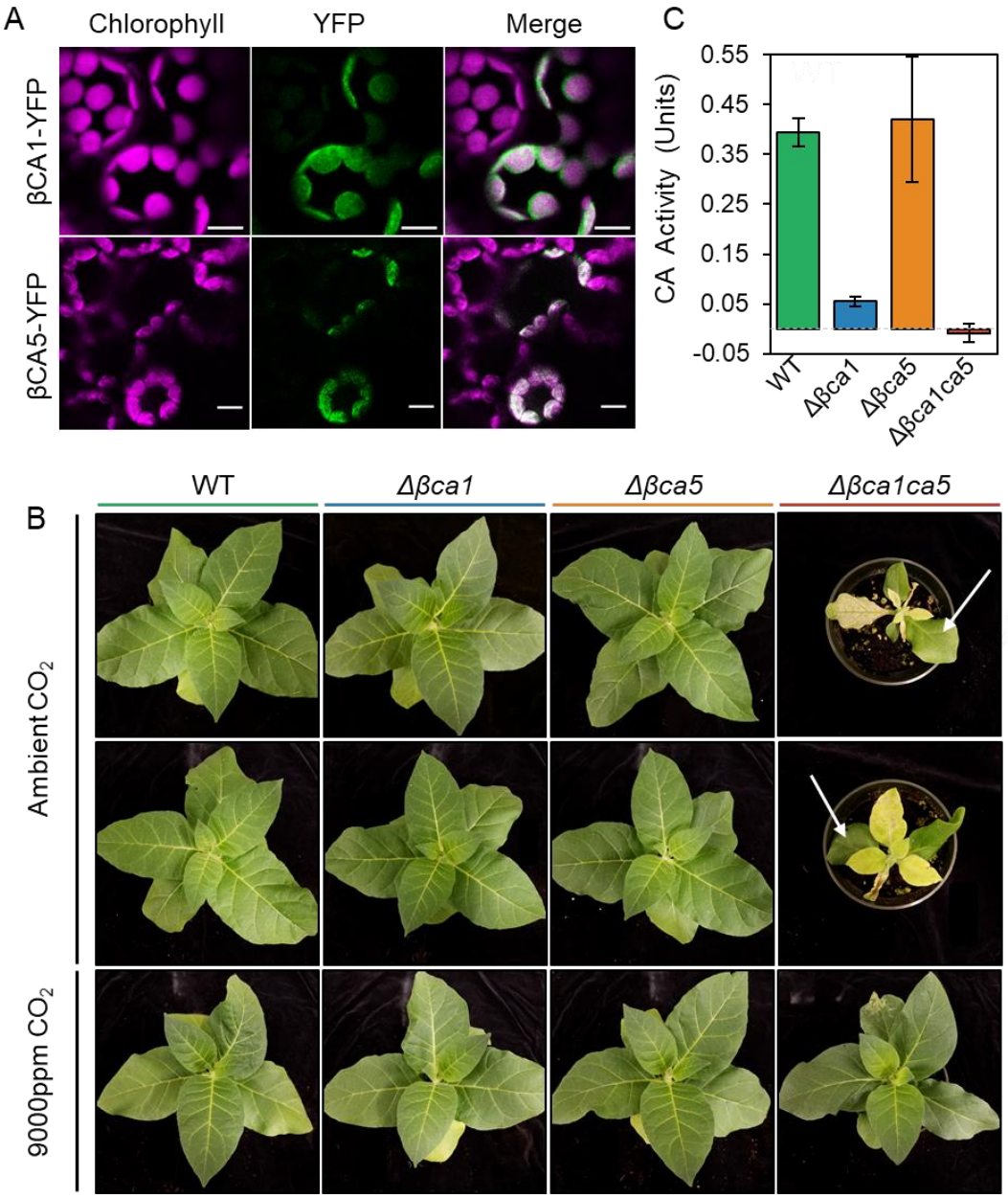

Figure 2.1. **Localization and CRISPR-generated mutations of tobacco chloroplast CAs**. **(A)** Confocal imaging of CA-YFP constructs transiently expressed in *N. tabacum* mesophyll cells. βCA1-YFP localized to the chloroplast stroma and cytoplasm. βCA5-YFP localized to the chloroplast stroma. Magenta: Chlorophyll autofluorescence, excitation 633nm. Green: YFP, excitation 514nm. Bars=10µm. **(B)** Single mutant lines *Δβca1* and *Δβca5* show no observable phenotype in their gross morphology compared to WT. The double mutant, *Δβca1ca5* produces pale leaves with symptoms of necrosis. Plants were grown in 9000ppm CO_2 for 4 weeks and then transferred to ambient CO_2 for 2 weeks before being imaged. White arrows indicate *Δβca1ca5* leaves that developed in high CO_2. **(C)** CA activity of whole leaf homogenate at ambient CO_2 (Bars: standard error).

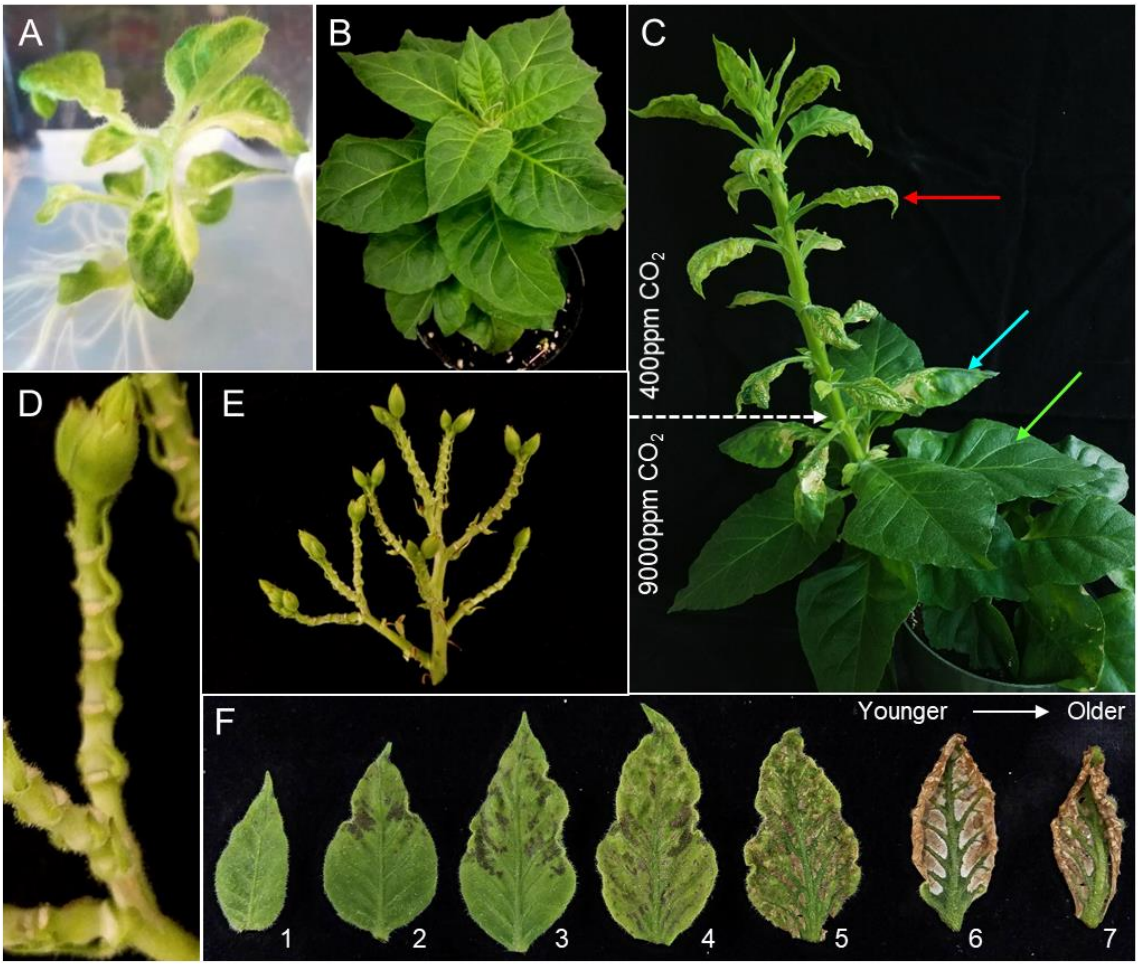

Figure 2.2. *Δβca1ca5* mutants have a severe developmental phenotype rescued by high CO_2. (A) *Δβca1ca5* shoot at ambient CO_2 growing on 3% sucrose media. **(B)** *Δβca1ca5* transplanted to soil and grown at 9000ppm CO_2. **(C)** *Δβca1ca5* mutant allowed to mature for 4 weeks at 9000ppm CO_2 before being transferred to ambient CO_2 (at white arrow) and allowed to grow for 4 more weeks. Green arrow: Example of a leaf that fully expanded at high CO_2. Red arrow: leaf that budded and developed at ambient CO_2. **(D&E)** Abscission of flowering buds in *Δβca1ca5* mutants at ambient CO_2. **(F)** Development of mutant phenotype in leaves developing at ambient CO_2 (left to right: youngest to oldest). Numbers indicate node position from top of plant.

Morphological characteristics of *Δβca1ca5* transgenic tobacco

The gross morphological and developmental phenotypes of the *Δβca1ca5* double mutant were observed at ambient CO_2 concentrations of about 415ppm (Figures 2.1B and 2.2A). When *Δβca1ca5* mutants were transplanted into soil and grown in 9000ppm CO_2, their morphology (leaf color and form) appeared to mirror WT tobacco (Figure 2.2B). Remarkably, leaves of *Δβca1ca5* that were able to fully expand at high CO_2 retained their WT-like morphology when they were transferred back to ambient CO_2 concentrations (Figure 2.2C, green arrow). These leaves did not become necrotic and remained green as the plant transitioned into flowering. Leaves that did not finish expanding in high CO_2 displayed necrotic lesions in their still-developing photosynthetic tissues (Figure 2.2C, blue arrow) and leaves that had initiated their development in ambient CO_2 also showed the mutant phenotype (Figure 2.2C, red arrow). When first emerging at ambient CO_2 concentrations, *Δβca1ca5* leaves do not show any obvious phenotypes or gross morphological differences from WT (Figure 2.2F, leaf 1). As these leaves age, however, necrotic lesions form and spread throughout the photosynthetic tissue. The midrib and primary veins do not show a cell-death phenotype (Figure 2.2F, leaves 6 and 7) and despite the death of mesophyll cells, the leaves remain attached to the stem of the tobacco plants. The pattern of necrosis (appearing first at the tip and then spreading to the base) matches the

pattern in which leaves transition from sink-to-source tissue (tip to base). This pattern is consistent with the lack of necrosis in sink tissues like the midrib and primary veins.

At ambient CO_2, *Δβca1ca5* plants also exhibit a unique flowering phenotype. Before the emergence of petals, the developing flower buds experience an early termination event, separating from the stem at the abscission zone (Figure 2.2D and 2E). A small number of buds eventually develop into mature flowers at ambient CO_2 and produce seeds via natural self-pollination.

CA mutants show no difference in photosynthetic capacity

The high content of CAs in the chloroplasts of C_3 plants (15) suggests that CA could play an important role in the availability of dissolved CO_2 for Rubisco. Such a role is likely to be most pronounced when demand for CO_2 is highest, during photosynthesis at saturating light. To investigate this, we imaged chlorophyll fluorescence *in vivo* from whole plants as leaves made the transition from limiting (7 µmol photons m^{-2} s^{-1}) to saturating (790 µmol photons m^{-2} s^{-1}) light. Φ_{PSII}, the efficiency with which absorbed photons are used to drive PS II electron transport, was measured at one-minute intervals beginning 10 seconds after the transition to saturating light. For accurate comparison, all plants were grown at 9000ppm CO_2 before being transferred to

ambient CO_2 for 24 hours for subsequent measurements. We observed that the kinetics of changes in Φ_{PSII} following the transition from limiting to saturating light (Figure 2.3A) were similar in WT and the mutants, as was the maximum value of Φ_{PSII} as the samples approached steady state photosynthesis at light saturation, indicating that photosynthesis was not impaired in the mutants compared to WT. Measurements of dark-adapted Fv/Fm (not shown) were all in the range of 0.75 to 0.81, suggesting that there is no accumulated photoinhibitory damage during growth in high CO_2 in either WT or mutant plants. Our photosynthetic measurements performed in ambient CO_2 indicate that enough CO_2 can diffuse into *Δβca1ca5* leaves for Rubisco to use without relying on converting the bicarbonate present in the stroma into CO_2.

ROS production is increased in *Δβca1ca5* leaves

Although increased ROS production is not predicted from the photosynthetic measurements, we assayed the level of ROS in WT and mutant lines (Figure 2.3B). Previous studies have shown plant CAs to play a role in preventing the creation of ROS (9, 13). Leaves from the first node of tobacco plants grown in ambient CO_2 were homogenized and the homogenate was suspended in a 5µM CellROX® Green solution. CellROX® Green undergoes a conformational change when oxidized, resulting in increased fluorescence. *Δβca1ca5* leaves displayed a greatly increased fluorescent signal in comparison to leaves of

Δβca1, Δβca5, and WT (Figure 2.3B). There was also a slight increase in the fluorescent signal in *Δβca5* over *Δβca1* and WT. This fluorescence pattern was also observed in the leaves of node 2 and 3 (data not shown). The location of the ROS signals was determined by injecting the 5µM CellROX® Green solution into mature leaves using a needleless syringe (the mature leaves had fully developed on plants grown in 9000ppm CO_2 and were then transferred to ambient CO_2 for 1 week). The ROS signals found in *Δβca1ca5* mesophyll cells appear to be associated with chloroplasts (Figure 2.3C), indicating the origins of the increased ROS measured in *Δβca1ca5* (Figure 2.3B).

Δβca1 and *Δβca1ca5* mutants show decreased immune response

As previously reported in Slaymaker, *et al.* (9), βCA1 is a salicylic acid binding protein that plays a role in the hypersensitive defense response of tobacco. Mature leaves of WT and CA mutant lines (all lines are Samsun-NN) were treated with 5µM of suspended tobacco mosaic virus (TMV) and grown at ambient CO_2. After three days, WT and *Δβca5* plants displayed the characteristic necrosis at the site of infection, indicating a hypersensitive response (Figure 2.3D). As expected, due to the lack of βCA1 enzyme present, neither *Δβca1* nor *Δβca1ca5* produced a hypersensitive response when inoculated with TMV.

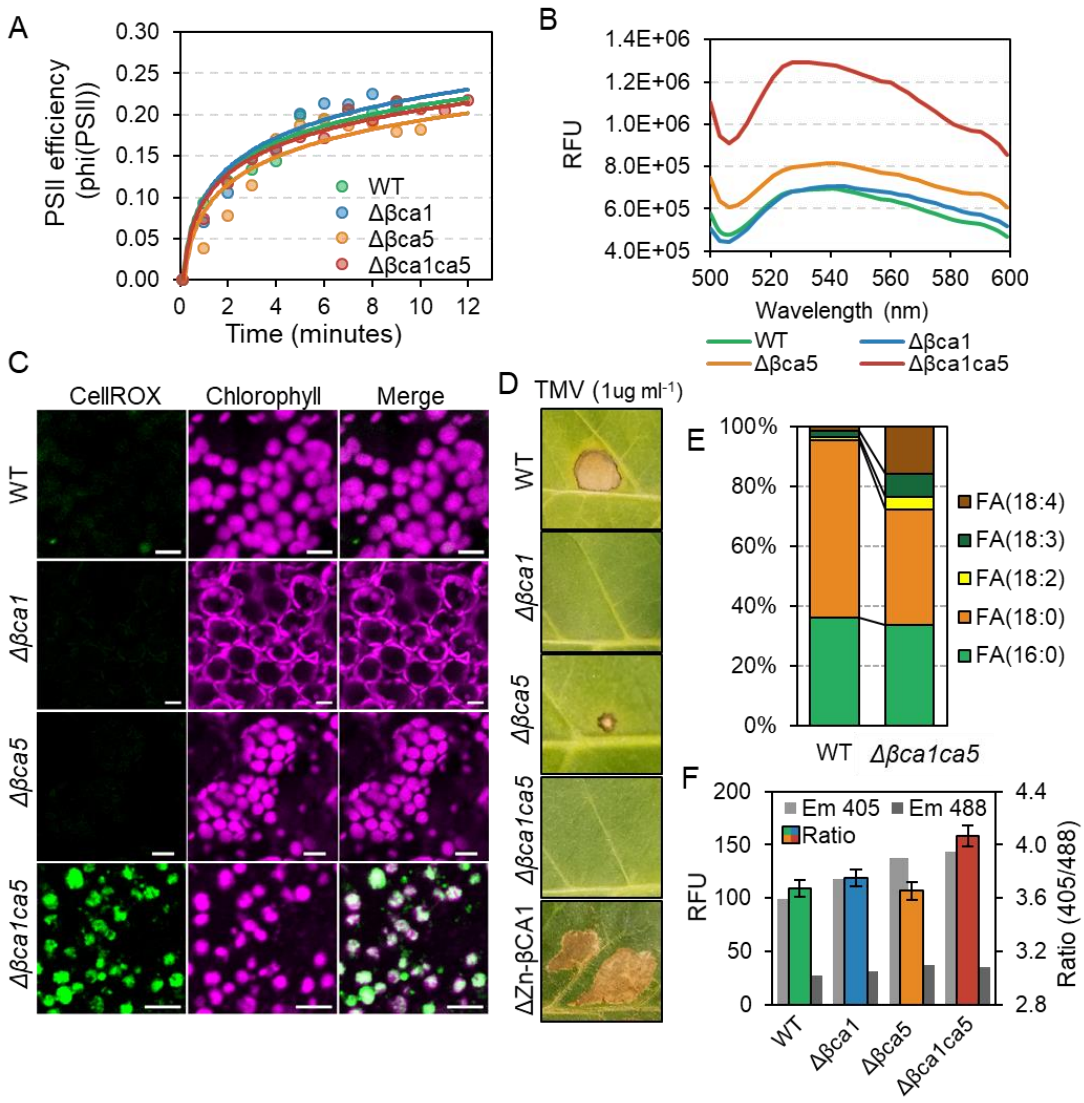

Figure 2.3. Analysis of mutant lines. (A) Light-response curve of WT tobacco and CA mutant lines with logarithmic trendlines plotted (R^2= 0.953, 0.924, 0.860, and 0.969 for WT, *Δβca1*, *Δβca5* and *Δβca1ca5* respectively). X-axis is minutes after transitioning to a saturating light intensity. **(B)** Spectrofluorometer emission spectra of CellROX® Green treated homogenized leaf samples. Signal intensity is in relative fluorescent units (RFU). **(C)** Confocal imaging of WT and mutant leaves grown in ambient CO_2 and treated with 5µM CellROX® Green. Magenta=Chlorophyll autofluorescence, Green=CellROX® Green, Bars=10µm. **(D)** TMV inoculation on fully expanded leaves of WT and transgenic lines. **(E)** Free Fatty Acid (FFA) analysis on ambient CO_2-grown leaves of WT tobacco and the Δ*βca1ca5* mutant line ($p \leq 0.001$) **(F)** Ratio (colored bars) of RecAcTP-pHlourin2 signals excited by 405nm (light grey) and 488nm (dark grey) lasers measured on a confocal microscope. Emission signals measured in relative fluorescent units ranged 0-255. Bars=Standard error.

Free fatty acid distribution is different between WT and *Δβca1ca5* leaves

Previously, Hoang and Chapman (24) used βCA1 antisense tobacco to demonstrate the important role CA plays in the incorporation of acetate into total lipids of chloroplasts. Here we investigated the free fatty acid (FFA) makeup of WT and Δ*βca1ca5* leaves (three young leaves from each the first three nodes) grown on sucrose media in ambient CO_2. A total of 5 FFAs were detected and annotated with this analysis (Figure 2.3E and Supplemental Figure S2.5). A significant difference between the two lines was observed. Saturated FFAs 18:2, 18:3, and 18:4 were upregulated in Δβ*ca1ca5* leaves whereas unsaturated FFAs 16:0 and 18:0 were higher in WT leaves (nomenclature for FFA is *number of carbons:number of saturated bonds*).

pH is increased in *Δβca1ca5* chloroplasts

In humans, CA plays a crucial role in the blood's bicarbonate buffer system (25), and while plant CAs have been hypothesized to play a role in regulating the pH of the chloroplast stroma (26), the level of bicarbonate at equilibrium in the stroma might be too low (~1 mM at equilibrium) for it to act as an effective buffer system (27). To investigate possible changes in pH in the CA mutant lines, the pH sensitive GFP pHluorin2 was altered with the addition of a chloroplast transit peptide (from the Arabidopsis RecA protein) to its N-terminus. This construct

was then transiently expressed in WT and mutant CA plants grown at ambient CO_2. While the exact pH cannot be determined with this method, it can give insights into the relative stromal pH of the mutants compared to the WT, which is expected to have a stromal pH of ~8 (28).

The ratio between the emission signals of pHluorin2 excited with the 405nm laser and the 488nm laser (405/488) was used to determine the relative pH of the cellular compartment (29). A higher 405nm/488nm ratio observed in WT chloroplasts indicates a basic environment, which is consistent with the predicted WT stromal pH of about 8 (Figure 2.3F). The pH of the single mutant lines *Δβca1* and *Δβca5* were similar to WT (p=0.11 and p=0.23 respectively. T-test) but *Δβca1ca5* showed a significant increase in the 405/488 ratio compared to WT (p=0.0001), indicating a higher pH and more basic cellular environment in the chloroplast stroma of the mutant than wild-type (Figure 2.3F).

Seed germination is severely reduced in Δβca1ca5

The T1 seed germination rates of the two single knockout lines, *Δβca1* and *Δβca5*, and the double mutant *Δβca1ca5* were measured in soil under ambient CO_2 concentrations and a long day (16h) photoperiod (Figure 2.4A). After 7 days, 97% of *Δβca1* seeds and 100% of WT seeds germinated. Germination of seeds from *Δβca5* was delayed, with 74% of seeds germinating by day 7, and 90%by

day 13. Only 3% of *Δβca1ca5* seeds germinated after 7 days; 15 days were required for the seeds to reach a germination rate of 29%.

The T1 *Δβca1* and *Δβca5* seedlings continued to develop normally under ambient CO_2 after germination (like their T0 parents). In contrast, the growth and development of T1 *Δβca1ca5* seedlings in ambient CO_2 arrested shortly after germination (Figure 2.4B). Notably, the *Δβca1ca5* seedlings in Figure 2.4B failed to produce true leaves at ambient CO_2, even when grown on media containing 3% sucrose.

The *Δβca1ca5* seeds shown in Figures 2.4A and 2.4B were produced from a T0 *Δβca1ca5* plant that flowered under 415ppm CO_2. We investigated whether growing *Δβca1ca5* plants in 9000ppm CO_2 would rescue the mutant phenotype observed in the seeds. The morphological phenotype of seeds from high CO_2-grown *Δβca1ca5* plants is similar to WT (Figure 2.4C). In comparison, seeds produced by *Δβca1ca5* in ambient CO_2 had significant morphological alterations and often contained large indentations in their coats (Figure 2.4C). The high CO_2 *Δβca1ca5* plants produced heavier seeds, averaging 0.068mg/seed compared to a mass of 0.038mg/seed produced from plants in ambient CO_2 (Figure 2.4D). In both atmospheric conditions, however, the mutant seeds weighed less than WT seeds ($p=0.00002$ and $p=0.0007$ for low and high CO_2 respectively, T-test),

which had an average mass of 0.081mg. To observe whether both mass and seed size differed, the two-dimensional areas of WT and *Δβca1ca5* seeds were measured. The seeds from ambient CO_2-grown mutants had an average area of 0.29mm^2, which was a statistically significant difference compared to the WT seed area of 0.32mm2 (Figure 2.4E, p= 0.0002, T-test). There was no significant difference in the average area of WT seeds 0.32mm^2 compared to the seeds of 9000ppm CO_2 grown *Δβca1ca5* plants, at 0.32mm^2 and 0.31mm^2 respectively (Figure 2.4E, p=0.52).

Seeds from *Δβca1ca5* plants produced in high CO_2 resulted in substantial improvement in seed germination rates, regardless of the CO_2 concentration in which the seeds were sowed. When grown on soil, these seeds had a 72% germination rate at ambient CO_2 and a 67% germination rate at high CO_2 (Figure 2.4F). Seeds produced by the *Δβca1ca5* plant grown at ambient CO_2 had germination rates of 29% and 22% when sown in soil and placed in 415ppm CO_2 and 9000ppm CO_2 concentrations, respectively. When placed on 3% sucrose media at 415ppm CO_2, the seeds from low CO_2-grown *Δβca1ca5* plants had a germination rate of 28% (Figure 2.4F). Thus, the growth conditions of the T0 *Δβca1ca5* parent had a strong influence on the germination rate of the ensuing seeds.

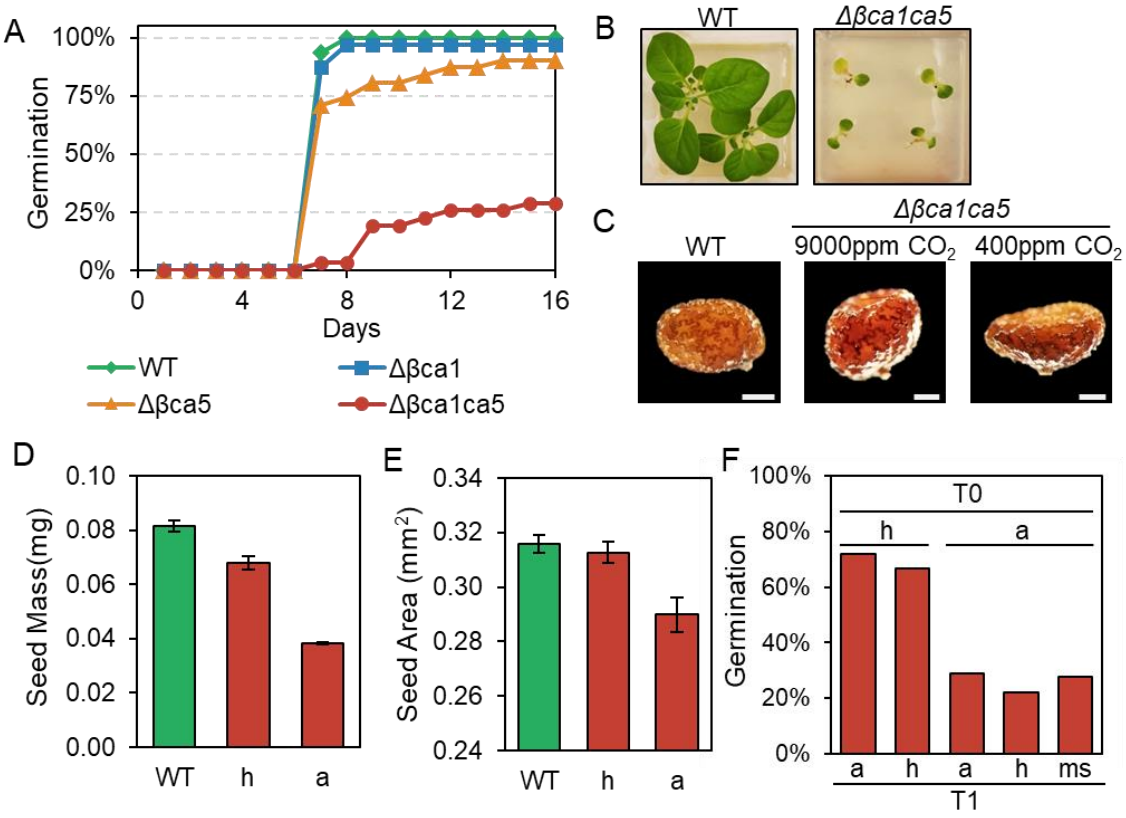

Figure 2.4. Analysis of CA mutant seeds. **(A)** Germination rates of WT seeds and T1 *Δβca1, Δβca5, and Δβca1ca5* seeds in soil at 400ppm CO_2. **(B)** WT and T1 *Δβca1ca5* seedlings, of the same age post sowing, on 3% sucrose media in 400ppm CO_2. **(C)** Sample comparison of T1 *Δβca1ca5* seeds with different growth conditions of the T0 mutant flowers. Bars= 0.2mm. (For D-F: h=9000ppm CO_2, a=400ppm CO2, ms=sucrose media) **(D)** Average mass of seeds from WT and *Δβca1ca5* plants grown under high or ambient CO_2 concentrations (bars= standard deviation). **(E)** Average size of seeds produced by WT and *Δβca1ca5* plants grown under high or ambient CO_2 concentrations (bars=standard error). **(F)** Effect of *Δβca1ca5* T0 plants growth condition on the germination rate of their T1 seeds.

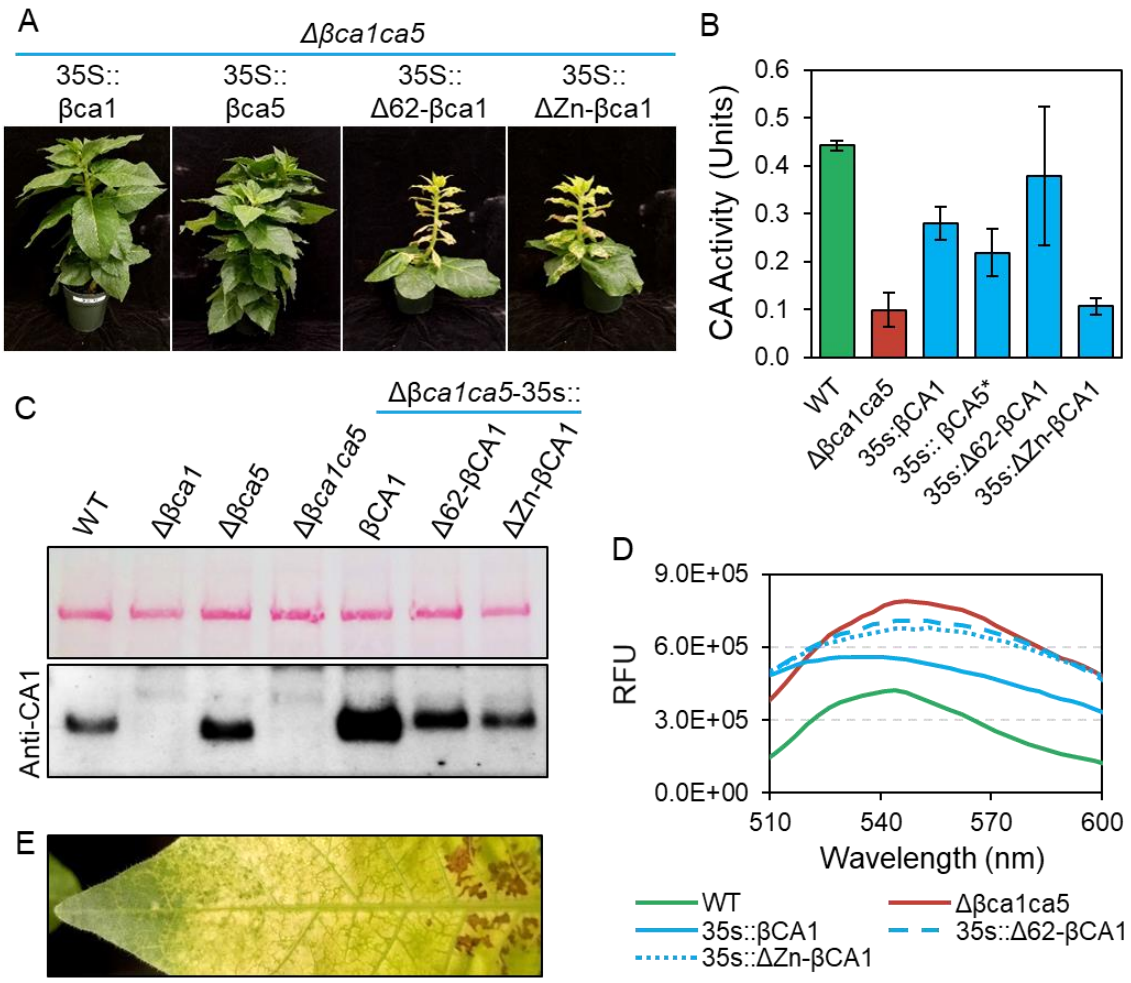

Figure 2.5. Complementation of CRISPR-generated *Δβca1ca5* mutant with different forms of CA proteins. (A) Complementation experiments performed in the *Δβca1ca5* transgenic line. **(B)** CA activity assay of WT, *Δβca1ca5*, and complemented lines (*No seeds produced, CA assay done on the T_0 generation). Bars=Standard error. **(C)** Top: Ponceau stain showing the rubisco large subunit signal. Bottom: Immunoblot of WT, mutant, and complemented lines leaf protein extracts with anti-CA1 antibody. **(D)** Spectrofluorometer emission spectra of CellROX® Green treated homogenized leaf samples. Signal intensity is in relative fluorescent units (RFU). **(E)** An expanding high-CO_2-grown *Δβca1ca5* leaf after being transferred to ambient CO_2.

Complementation of Δβ*ca1ca5* mutants

To confirm that the high CO_2 phenotypes observed in *Δβca1ca5* mutants were not a result of off-target mutations caused by the CRISPR/Cas9 system, a series of complementation experiments were performed. For these experiments, *Δβca1ca5* T1 plants lacking the *Cas9* transgene were identified (Supplemental Figure S2.4). Transgene constructs in which the βCA1 CDS and βCA5 CDS were put under the control of the 35S promoter were transformed into *cas9⁻ Δβca1ca5* T1 plants (Supplemental Figure S2.6). Both *Δβca1ca5*_35S::βCA1 and *Δβca1ca5*_35S::βCA5 complemented the mutant phenotype and produced plants with a gross morphology comparable to WT at ambient CO_2 concentrations (Figure 2.5A). This result establishes that the phenotype observed in *Δβca1ca5* tobacco is caused by lack of stromal CA activity and not by an off-target mutation.

The WT phenotype could not be restored to *cas9⁻Δβca1ca5* plants expressing altered forms of the βCA1 enzyme. We generated a construct in which two of the residues that bind zinc in βCA1 were mutated to adenine (C152A and D152A; shortened hereafter as ΔZn), in order to make a catalytically inactive form of βCA1 (Supplemental Figure S2.6). The zinc ion located at the active site cavity of βCAs is essential for the nucleophilicity of the enzyme and thus, its catalytic

activity (30). Because our YFP-fusion data showed potential dual localization of the over-expressed full-length βCA1 CDS to both the cytoplasm and the chloroplast (Supplemental Figure S2.1A), we also generated a construct in which the chloroplast transit peptide (identified as the first 62 amino acids) was removed from βCA1. ΔZn-βCA1 and Δ62-βCA1 were both driven by a 35S overexpression promoter and transformed into *cas9⁻Δβca1ca5* plants. Neither of these constructs resulted in transgenic plants with normal morphology; instead the plants showed the characteristic, necrosis-like phenotype at ambient CO_2 (Figure 2.5A).

An immunoblot of extracted proteins using an anti-CA1 antibody showed that the mutant and full-length forms of βCA1 were present in the three transgenic lines used in the complementation experiments (Figure 2.5C), indicated by a positive band at ~28kDA in leaf homogenate of WT, Δ*βca5*, *Δβca1ca5*_35S::βCA1, *Δβca1ca5*_35S::Δ62-βCA1, and *Δβca1ca5*_35S::ΔZn-βCA1 tobacco plants.

CA activity assays determined that *cas9⁻Δβca1ca5* plants expressing the cytoplasm-localized βCA1 (Δ62-βCA1) and full-length βCA1 or βCA5 had substantially increased CA activity compared to the double mutant (Figure 2.5B), signifying that Δ62-βCA1 is still catalytically active. Both *Δβca1ca5*_35S::Δ62-

βCA1 and *Δβca1ca5*_35S::ΔZn-βCA1 had heightened ROS levels like *Δβca1ca5*, while *Δβca1ca5*_35S::βCA1 showed a decrease in ROS production (Figure 2.5D). When inoculated with TMV (as previously described), *Δβca1ca5*_35S::ΔZn-βCA1 plants display the characteristic programed cell death of an HR, despite their minimal CA activity levels (Figure 2.5B). This result suggests that ΔZn-βCA1 is still able to bind to SA and contribute to plant defense even with a mutated catalytic center.

Taken together, these experiments indicate that both the location of βCA1 (in the chloroplast stroma) and its catalytic activity are necessary to complement the *Δβca1ca5* phenotype, except for the altered hypersensitive response.

DISCUSSION

Previous experiments that studied single knockouts of βCA1 did not eliminate all enzymatic activity from the chloroplast stroma because no lines with both stromal CAs mutated were examined (9, 16, 27). As a result of eliminating expression of both chloroplast stromal CAs, we have observed a striking defect in leaf development as well as impaired seed formation. Any potential explanations for the dramatic, CO_2-sensitive phenotype in the double mutant need to address the following two observations: (1) the CO_2-dependent phenotype (i.e. the WT-like

growth of leaves at high CO_2 concentrations) and (2) the specificity of the leaf phenotype, which affects leaves only before they are fully expanded.

Our photosynthetic measurements suggest that enough CO_2 can diffuse into *Δβca1ca5* leaves for Rubisco to use without relying on converting the bicarbonate present in the stroma into CO_2. The lack of deficiencies in photosynthesis matches observations by Price, *et al.* (16) where antisense-CA1 mutant tobacco, with a 98% reduction in CA activity, showed no measurable reduction in Rubisco activity levels nor CO_2 uptake. There was, however, a small, but not statistically significant increase in stomatal conductance in antisense-CA1 lines that correlated with a reduction in CA activity (16). For normal plant development, evidently 2% of CA is adequate, likely due to its high rate of catalysis.

While a defect in photosynthesis does not appear to be the cause of the phenotype observed in *Δβca1ca5*, the surge in chloroplast ROS could still be a potential explanation for the leaf phenotype, albeit necessarily arising from a pathway other than photosynthesis. A possible hypothesis for the necrotic *Δβca1ca5* phenotype is that programed cell death is somehow being triggered by the complete lack of stromal CA, which might classify the double mutant as a lesion mimic (31).

βCA1 has been shown to participate in the plant defense system, given that it can bind to the plant hormone salicylic acid (SA) (9). When Slaymaker, *et al.* (9) silenced βCA1 in *Nicotiana tabacum*, they observed an absence of a HR when the leaves were transiently transformed by Agrobacterium expressing *Pto* and *avrPto*. We also observed a lack of HR in *Δβca1* and *Δβca1ca5* plants (Figure 2.3D).

βCA1 binds SA and is involved in the perception of SA levels in plants (10, 32). *Δβca1ca5* plants expressing a zinc-binding site mutated βCA1 (ΔZn-βCA1) were unable to complement the developmental mutant phenotype at ambient CO_2 concentrations (Figure 2.5A), even though the inactive protein was present in leaf extracts (Figure 2.5C). Δβ*ca1ca5*_35S::ΔZn-βCA1 plants showed a hypersensitive response when inoculated with TMV, which implies that its ability to bind SA has not been compromised by the zinc-binding site mutations at its catalytic core (Figure 2.3D). These plants can induce programmed cell death in response to a pathogen as WT tobacco can, but still demonstrate leaf necrosis at ambient CO_2, which reflects the fact that the ΔZn-βCA1 has no enzymatic activity.

As evident from considering Equation 1.1 (Chapter 1), even though CA is not required for photosynthesis, at ambient CO_2 there may not be enough bicarbonate

for other types of reactions in which substrates are carboxylated. The specific reactions that may be affected are discussed in an excellent review by DiMario, Machingura, Waldrop and Moroney (27). Without CA to replenish bicarbonate, several vital pathways could be starved, which would explain the drastic phenotype we observed. Recovery of the double mutant in high CO_2 can be explained by shifting of Equation 1.1 (Chapter 1) to the right. (The rate of the uncatalyzed conversion of CO_2 is increased when the concentration of CO_2 is increased). The lack of bicarbonate will have significant effect on pathways with enzymes that have a low affinity for bicarbonate, like 5-aminoimidazole ribonucleotide carboxylase in the purine synthesis pathway, which has Km (mM) of 23 (33), and pathways that use a lot of bicarbonate, like fatty acid synthesis (34-36). The latter is discussed below.

A fascinating aspect of the double CA mutant is its defect in leaf development. *Δβca1ca5* leaves will not display the mutant phenotype if allowed to fully expand under high (9000ppm) CO_2 (Figure 2.2C, green arrow). Leaves do not develop uniformly: In developing leaves, the sink-to-source transition happens first at the leaf tip and ends at the base. When developing *Δβca1ca5* leaves are moved from high CO_2 to ambient CO_2, areas of the leaf that have likely completed a transition to source tissue (the tip) do not display the mutant phenotype, while areas that have not completed transition to a sink (the base) develop the characteristic

lesions (Figure 2.2C, blue arrow and Figure 2.5E). Interestingly, leaves which bud and develop completely at ambient CO_2 show a different pattern where necrosis begins at the tip and spreads to the base (Figure 2.2F).

The phenotype on the transition from high to ambient CO_2 fits closely with what is observed in the sink-to-source transition of developing leaves. The tips of the leaves developing in ambient CO_2 in Figure 2.2F are the first cells to undergo a sink-to-source transition and they are also the first to exhibit signs of the mutant phenotype. In Figure 2.5E, the cells near the tip likely have completed their sink-to-source transition and appear green and healthy whereas tissue near the base is dying. However, although newly budded leaves of the double CA mutant in ambient CO2 appear like WT in gross morphology, the entire leaf does exhibit higher ROS localized to the chloroplast (Figures 2.3B and 2.3C), which may set the pathway to necrosis in motion.

Molecules whose synthesis could be affected by inadequate bicarbonate (and whose absence might contribute to the double mutant phenotype) include malonate and its derivatives, which are used in the synthesis of lipids (37), and flavonoid derivatives that are critical for plant defenses against pathogens and UV light (38, 39). Malonate and fatty acid synthesis depends on the enzyme Acetyl-CoA Carboxylase (K_m 0.9-2.5mm), which uses bicarbonate to

carboxylate its substrate (27, 34-36). When CA activity in tobacco leaves is suppressed using inhibitors (like ethoxyzolamide) or antisense RNA silencing, the incorporation of acetate into total lipids is greatly reduced (24). We observed a significant difference between the FFA profiles of young WT and *Δβca1ca5* leaves (Figure 2.3E). The double mutant had a significant reduction in the amount of 16:0 and 18:0 FFA, both of which arise in the chloroplast before being elongated and further modified in other organelles. This data is consistent with a reduction of the biosynthesis of FFA in the chloroplast in the double mutant compared to the WT at ambient CO_2 concentrations.

This hypothesis would also explain the changes in seed morphology and seedling survivability we observed. Suppressing CA activity results in improper incorporation of acetate into lipids of developing cotton (*Gossypium hirsutum)* embryos and an increase in seedling lethality in Arabidopsis (24, 40). Importantly, Hoang and Chapman (24) demonstrated that it was the inhibition of plastid CA activity that resulted in the reduced incorporation of acetate into lipids in cotton embryos. This explanation is likely the reason why the seeds of the double CA tobacco mutant were unhealthy and had lower germination rates when grown in ambient CO_2.

The *Δβca1ca5* mutant showed an increased ratio between the 405nm and 488nm emission signals, indicating a more basic environment compared to WT and the single mutants *Δβca1* and Δ*βca5*. It has been postulated that CA activity could contribute to a bicarbonate buffering system that mitigates pH-transients from the photosynthesis light reactions. While the pH measurements we performed were not quantitative, they provide an interesting look into how bicarbonate and carbon dioxide are utilized in the chloroplast without CA present to quickly re-establish an equilibrium.

Without CA, the pools of CO_2 and bicarbonate in the stroma are essentially isolated under light conditions (where CO_2 is constantly being used in photosynthesis), likely to reach equilibrium only in the dark. Since CO_2 can diffuse across the chloroplast envelope, photosynthesis is not dependent on the reverse hydration of bicarbonate to create enough CO_2 for its purposes. If the reverse reaction was required, and thus CO_2 was limiting, we would expect the stroma pH to decrease in the *Δβca1ca5* mutant as it becomes more acidic due to the now relatively higher levels of bicarbonate (that, with CA present, would be used to generate CO_2). Our mutants had no apparent complications in photosynthesis and instead we observed an increase in the stroma pH. This is the result one would expect if CO_2 was *not* limiting and instead bicarbonate, which is used in many pathways, was inadequate. As bicarbonate is used up, it would

be very slowly replaced by the uncatalyzed reaction and therefore the pH would increase due to the inability to quickly generate the bicarbonate and proton seen in Equation 1.1 (Chapter 1).

However, pH also plays a role in regulating the Calvin cycle (41) and our photosynthesis measurements described above give no indication of changes in Calvin cycle activity between WT and the double mutant. Another contribution to the disturbed developmental processes could be that critical biosynthetic enzymes might be unable to function properly in the presence of higher pH.

In summary, CRISPR-generated knockout lines in tobacco demonstrate that CA activity in the chloroplast stroma is required for normal plant and seed development at ambient CO_2 concentrations, but not for supplying CO_2 for photosynthesis. Removing CA activity from the chloroplast stroma results in large necrotic lesions in developing leaf tissue and reduced germination of seeds following growth in ambient CO_2. Both mutant phenotypes can be prevented by growing *Δβca1ca5* plants in high CO_2 concentrations of 9000ppm. The mechanism by which removing chloroplast CA results in necrotic lesions remains to be elucidated.

Given that the mutants lacking CA are normal at elevated CO_2, which results in increased bicarbonate concentration of the stroma, then transport of bicarbonate into chloroplasts in ambient CO_2should also be able to supply adequate bicarbonate for biosynthetic reactions. In cyanobacteria, where CA is confined carboxysomes, bicarbonate transporters on the plasma membrane supply bicarbonate to the cytoplasm. Schemes to produce a functional carboxysome in chloroplasts propose incorporation of bicarbonate transporters on the chloroplast envelope membrane (18, 42, 43). We have shown that the hypersensitive response could be maintained in plants engineered to lack CA, provide that an inactive CA is incorporated to provide virus resistance. Thus, our work supports the feasibility of an important step in the engineering a carbon-concentrating mechanism into a C_3 plant.

METHODS

YFP fusion localization

The TargetP online prediction program (44) was used identify candidate CA proteins that might be located in the chloroplast. RNA was isolated from mature *N. tabacum* plants and reverse transcription was performed using universal primers to obtain cDNA. βCA1, βCA2, βCA5, and αCA1 cDNAs were amplified and PCR products were inserted into the pCR8/GW/TOPO® vector by TA cloning. The CA sequences were then fused to YFP at their C-termini by insertion

into the pEXSG-eYFP vector via an LR clonase II reaction. Electrocompetent *Agrobacterium tumefaciens* (GV3101::RK) cells were transformed with the pEXSG-CA-eYFP vectors and then used to transiently transformed young *N. tabacum* plant leaves (45). The plants were imaged on a confocal microscope after 48 hours. Primers used for the creation of constructs can be found in table 2.1. The βCA5 cDNA sequence has been deposited onto GenBank with the accession number BankIt2241503 beta_carbonic_anhydrase_5 MN153507.

Design and assembly of CRISPR/Cas9 constructs

The CCTop online program was used to identify and select CRISPR/Cas9 sgRNA target sites in the βca1 and βca5 genes (46). Target sites were chosen to be <150bp apart to create large deletions that could easily be genotyped by PCR. Constructs containing the Cas9 enzyme gene and sgRNA sequences were assembled using the Golden Gate cloning method (47). For the single knockout lines, the two sgRNA sequences for either βca1 or βca5 were put under control of an Arabidopsis ubiquitin 6 promotor (from pICH86966) and inserted into pICH47751 and pICH47761 plasmids. The sgRNA constructs along with the Cas9 (pICH47742 2x35S-5'UTR-hCas9), the kanamycin selectable marker (pICSL11024 NOSp-NPTII-OCST), and the linker (pICH41780) constructs were then assembled into the final, level-2 construct pAGM4723 and transform into

plant tissue (48, 49). For the double knockout lines, pICH47781 and pICH47791 were used for the βca5 sgRNA targets.

Agrobacterium-mediated transformation of tobacco

Transformation of tobacco generally followed the protocol described in Sparkes et al. (2006). Briefly, leaves from two-month-old tobacco plants were injected with a liquid culture of GV3101 (*Agrobacterium tumefaciens*) cells measuring an OD_{600} of 0.5. Inoculated plants were grown in long day (16hr) chambers for two days. The inoculated leaves were sterilized in 75% ethanol for 10 seconds and 20% bleach solution for 20 minutes. MS104 (regular MS media with 6-Benzylaminopurine [Sigma B3274] and α-naphthalene acetic acid [Sigma N1641]) media containing 3% sucrose and 75μg/mL kanamycin was used for selection and generation of transformed shoots.

Genotyping and sequencing transformed lines

Rooted transformed plants were used in genotyping. With primers specific for Cas9, the MyTaq^{TM} Plant-PCR kit (BIO-25055) was used to identify plants containing the T-DNA insertion and to determine whether there was a change in the PCR product derived from the CA genes. Genomic DNA was isolated from transgenic plants with cetyl trimethylammonium bromide (CTAB). Phusion DNA polymerase was used to amplify the mutated CA genes. Because of the

possibility that there could be a variety of mutations present in PCR products of approximately the same size, a colony library was created by cloning the PCR products into the pCR®2.1 TOPO vector, transforming *E. coli*, and plating. Colonies were cultured, and the individual plasmids were isolated and sequenced.

CA activity assay

The CA activity assay was carried out as described in Medina-Puche, *et al.* (10) with slight modifications. Small leaf samples from mature plants were homogenized in 0.2M Tris-HCl, pH8.3, 1mM EDTA, 20mM $MgCl_2$, 50mM NaCl, and 0.1M Na_2SO_4 in a 3:1 ratio (volume:weight). 100µL of homogenate was mixed with 200µL of 20mM Tris-HCl, pH8.3 with bromothymol blue. The reaction was started by adding 700µL of cold carbonated ddH_2O. Water was carbonated in a SodaStream® system (item model number 1012111015) and incubated on ice for 30 minutes before use in assays. The time it took for the mixture to drop from pH8.3 to 5.5 (indicated by the dye turning from blue to yellow) was recorded, with the same procedure being followed in a control tube lacking tissue. The units of enzyme activity were calculated as: *1 unit = (Tb/Ts-1)/(Ts*P)* , where *Tb* and *Te* represent the time (in seconds) it took for the pH to drop from 8.3 to 5.5 in the control and plant homogenate reactions, respectively (50), and represent the protein concentration in milligrams. Each sample and control were done in replicates of four and reactions took place on ice.

ROS Assay

Emergent leaves from the first node of 10-week-old tobacco plants grown for 8 weeks at 9000ppm CO_2 and 2 weeks at ambient CO_2 were used to measure the presence of ROS using the fluorescent dye CellROX® Green (Thermofisher Scientific Cat. No. C10444). The original protocol was adapted for use on plant tissues. Briefly, whole leaf samples were frozen in liquid nitrogen, grounded into a powder, and suspended in 5µM of CellROX® Green in homogenization buffer (0.2 M Tris-HCl, pH8.3, 1 mM EDTA, 20 mM $MgCl_2$, 50 mM NaCl, 0.1 M Na_2SO_4). After being incubated at room temperature for 30 minutes, samples were then measured on a PTI spectrofluorometer using an excitation wavelength of 480nm. For confocal imaging CellROX® Green was diluted in to 5µM in ddH_2O and injected into expanding leaves at the third node using a needleless syringe. Treated plants were incubated for 30 minutes at room temperature. Leaf samples were cut out and imaged on a Zeiss LSM710 inverted confocal microscope using the 488nm laser line.

pH measurements with pHlourin2

The plasmid pME-pHluorin2 was a gift from David Raible (Addgene plasmid # 73794 ; http://n2t.net/addgene:73794 ; RRID:Addgene_73794) (29, 51). The pHlourin2 protein sequence was altered with the addition of the chloroplast transit peptide from the Arabidopsis RecA DNA recombination family protein

(accession NM_001084375.1) to its N-terminus. The plasmids pBI121-gus +3xFLAG-Strep+GWb and pME-RecAcTP-pHluorin2 were used in a Gateway cloning reaction to produce pBI121-RecAcTP-pHluorin2. WT and transgenic tobacco lines were transiently transformed with Agrobacterium similarly to the *YFP fusion localization* methods described above. The plants grew in ambient CO_2 (16-hour days, 61µM light) and were imaged on a confocal microscope 38- and 72-hours post-transformation. The samples were excited under light conditions with 633nm (2.0%), 405nm (3.0%), and 488nm (17%) wavelength lasers with a detectable emission range between 500nm and 530nm (for the 405 and 488 laser lines). Emission signal intensity was measured in chloroplasts with the ratio being calculated by dividing the average signal intensity from the 405nm laser by the 488nm laser.

Chlorophyll fluorescence measurements

Chlorophyll fluorescence was measured on intact plants using WALZ Imaging PAM M-series chlorophyll fluorometer (Heinz Walz GmbH, Effeltrich, Germany). Samples were positioned such that the leaf surfaces were between 18 and 19 cm from the LED array. All samples were adapted to a light intensity of 7 µmoles photons m^{-2} s^{-1} for a minimum of 30 minutes before measurement. Samples were then imaged using the non-actinic measuring light (470 nm, frequency = 1Hz) and three to five 1-cm diameter areas of interest (AOI) were

defined on sections of each leaf that did not include major veins. Leaves selected for measurement were developmentally equivalent among treatments. After an initial measurement of the steady state fluorescence at 7 μmoles photons $m^{-2}s^{-1}$ actinic intensity, the plants were exposed to continuous 790 μmoles photons m^{-2} s^{-1} light from the LED array. Starting at 10 seconds after the increase in actinic intensity, Φ_{PSII} was measured at one-minute intervals using a saturating pulse intensity of 1790 μmoles photons m^{-2} s^{-1} until a steady state was reached (typically 8-12 minutes). Reported Φ_{PSII} values were averages over 3-5 AOIs per leaf, 2-3 leaves per plant and 2-4 plants per genotype.

Sample preparation for free fatty acid distribution

Young leaves from the first 3 nodes of WT tobacco and the *Δβca1ca5* mutant were ground in a liquid-nitrogen treated mortar and pestle to create 3 replicates of 0.25g fresh-weight powder. 100mg of sample were placed in a high force Eppendorf tube, and after adding 300mL of ultrapure water, the mixture was vortexed for 30s. 187μL of DCM/MeOH (2:1, by volume) was added, vortexed, and the samples were left on ice 30 min. After the 30 min, 0.4 mL of Dichloromethane (DCM) were added and vortexed and then another another 0.4 ml of ultrapure water were added and vortexed again. Finally, the extracts were centrifuge at 13,000 rpm for 15 minutes at 4 °C. The bottom organic layer was

collected into glass tubes and dried down by speed vacuum. The residues were reconstituted in 150µL of IPA/ACN/H2O (65/30/5) that contains Heptadecanoic acid as ISTD for normalization.

Samples were sent to the Metabolomics Facility at Cornell University's Biotechnology Resource Center. Analysis was conducted on a Thermo Scientific™ Vanquish™ Horizon UHPLC / Thermo Scientific™ Q Exactive™ HF hybrid quadrupole-Orbitrap mass spectrometer.

TMV Inoculation

Inoculation with TMV was performed similar to the method of Guo, Salih and Klessig (52). WT and mutant tobacco plants were grown at 9000ppm CO_2 for 8 weeks before being transferred to ambient CO_2. The genetic background of all lines is *Nicotiana tabacum* L. cv. Samsun-NN, which is resistant to TMV as a result of the *N* gene product interfering with the TMV replicase (53-55). Two fully expanded leaves were chosen from each plant for inoculation (n=10). TMV was diluted to 1µg mL^{-1} in 25 mM Sodium phosphate buffer pH 7.2. Carborundum was sprinkled on the leaves to induce wounding and allow virus entry. A cheesecloth was rinsed in the phosphate solution and then soaked in the diluted TMV/Sodium phosphate solution. Cheese cloth was rubbed on selected

leaves and plants were incubated for 45 minutes before being washed with clean ddH_2O. Plants were imaged 6 days post-inoculation.

Generation of complementation constructs

Full-length βCA1 and βCA5 CDS were generated in via RT-PCR for YFP fusion. Primers (which added SpeI and AscI RE sites) were used to re-add the stop codon to the full-length CDS and additional primers were used to generate Δ62-βCA1. The plasmid pK7WGF2::hCas9 (AddGene: 46965) was used as the backbone and a restriction/digestion was performed to replace the 35S-driven hCas9 gene with βCA1, βCA5, and Δ62-βCA1. The zinc-binding mutant was generated using NEB's Q5® Site-Directed Mutagenesis Kit and primers designed by the NEBaseChanger™ tool. *Δβca1ca5* Cas9$^-$ Plants were transformed and selected as described in the "Agrobacterium-mediated transformation of tobacco" section. Plants were genotyped using primers that flanked of insert on the plasmid sequence.

Growth chamber conditions

Plants were grown under 800 μmol/m^2/s photosynthetically active radiation (PAR) with a 16-hour photoperiod in ambient and high (9000ppm) CO_2 growth chambers. Growth room temperatures were maintained at an average of 22^{o}C with 50% relative humidity during daylight conditions for the ambient CO_2

chamber and 64% relative humidity for the high CO_2 chamber. Plants were germinated and grown in pots containing LM-111 All Purpose Mix (Lambert), watered daily in the morning by hand, and kept in free-draining flats.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Joyce Van Eck and Dr. Daniel Klessig of the Boyce Thompson Institute for the plasmids containing Cas9+sgRNA and for use of their growth facilities for the TMV inoculation, respectively. We would also like to thank Dr. M. Elena Diaz Rubio (Cornell Metabolomics Facility) for performing mass spectrometry for fatty acid composition. K.N.E. participated in the Cornell Department of Molecular Biology and Genetics NSF REU program. The work was partially supported through a Cornell University Institute of Biotechnology's Center for Advanced Technology (CAT) grant, funded through New York State Division of Science, Technology, and Innovation (NYSTAR) and NYS contract C150124 (grant number NIH S10RR025502 for data collected on the Zeiss LSM 710 Confocal). Major funding was from Bilateral NSF-BIO/BBSRC 1642386 to M.R.H.

REFERENCES

1. D. Hewett-Emmett, R. E. Tashian, Functional diversity, conservation, and convergence in the evolution of the alpha-, beta-, and gamma-carbonic anhydrase gene families. *Mol Phylogenet Evol* **5**, 50-77 (1996).
2. X. Ren, S. Lindskog, Buffer dependence of CO_2 hydration catalyzed by human carbonic anhydrase I. *Biochim Biophys Acta* **1120**, 81-86 (1992).
3. J. V. Moroney, S. G. Bartlett, G. Samuelsson, Carbonic anhydrases in plants and algae. *Plant, Cell and Environment* **24**, 141-153 (2001).
4. M. D. Hatch, J. N. Burnell, Carbonic anhydrase activity in leaves and its role in the first step of c(4) photosynthesis. *Plant Physiol* **93**, 825-828 (1990).
5. M. R. Badger, G. D. Price, The role of carbonic anhydrase in photosynthesis. *Annual Review of Plant Physiology and Plant Molecular Biology* **45**, 369-392 (1994).
6. P. H. Su, Y. H. Lai, A reliable and non-destructive method for monitoring the stromal pH in isolated chloroplasts using a fluorescent pH probe. *Frontiers in plant science* **8**, 2079 (2017).
7. N. Fabre, I. M. Reiter, N. Becuwe-Linka, B. Genty, D. Rumeau, Characterization and expression analysis of genes encoding alpha and beta carbonic anhydrases in Arabidopsis. *Plant Cell Environ* **30**, 617-629 (2007).
8. R. J. DiMario *et al.*, The cytoplasmic carbonic anhydrases betaCA2 and betaCA4 are required for optimal plant growth at low CO_2. *Plant Physiol* **171**, 280-293 (2016).
9. D. H. Slaymaker *et al.*, The tobacco salicylic acid-binding protein 3 (SABP3) is the chloroplast carbonic anhydrase, which exhibits antioxidant activity and plays a role in the hypersensitive defense response. *Proc Natl Acad Sci U S A* **99**, 11640-11645 (2002).
10. L. Medina-Puche *et al.*, beta-carbonic anhydrases play a role in salicylic acid perception in Arabidopsis. *PLoS One* **12**, e0181820 (2017).
11. H. Hu *et al.*, Distinct Cellular Locations of Carbonic Anhydrases Mediate Carbon Dioxide Control of Stomatal Movements. *Plant Physiol* **169**, 1168-1178 (2015).
12. H. Hu *et al.*, Carbonic anhydrases are upstream regulators of CO_2-controlled stomatal movements in guard cells. *Nature Cell Biology* **12**, 87-93; sup pp 81-18 (2010).

13. J. Dabrowska-Bronk *et al.*, beta-carbonic anhydrases and carbonic ions uptake positively influence Arabidopsis photosynthesis, oxidative stress tolerance and growth in light dependent manner. *J Plant Physiol* **203**, 44-54 (2016).

14. C. V. Hoang, K. D. Chapman, Biochemical and molecular inhibition of plastidial carbonic anhydrase reduces the incorporation of acetate into lipids in cotton embryos and tobacco cell suspensions and leaves. *Plant Physiol* **128**, 1417-1427 (2002).

15. K. Okabe, S.-Y. Yang, M. Tsuzuki, S. Miyachi, Carbonic anhydrase: Its content in spinach leaves and its taxonomic diversity studied with anti-spinach leaf carbonic anhydrase antibody. *Plant Science Letters* **33**, 145-153 (1984).

16. G. D. Price *et al.*, Specific reduction of chloroplast carbonic anhydrase activity by antisense RNA in transgenic tobacco plants has a minor effect on photosynthetic CO2 assimilation. *Planta* **193**, 331-340 (1994).

17. A. Pal, D. Borthakur, Transgenic overexpression ofLeucaenaβ-carbonic anhydrases in tobacco does not affect carbon assimilation and overall biomass. *Plant Biosystems* 150, 932-941 (2015).

18. M. R. Hanson, M. T. Lin, A. E. Carmo-Silva, M. A. J. Parry, Towards engineering carboxysomes into C3 plants. *The Plant Journal* **87**, 38-50 (2016).

19. N. Atkinson *et al.*, Introducing an algal carbon-concentrating mechanism into higher plants: location and incorporation of key components. *Plant biotechnology journal* **14**, 1302-1315 (2016).

20. M. R. Badger, G. D. Price, CO_2 concentrating mechanisms in cyanobacteria: molecular components, their diversity and evolution. *Journal of experimental botany* **54**, 609-622 (2003).

21. G. D. Price, M. R. Badger, Expression of human carbonic anhydrase in the cyanobacterium *Synechococcus* PCC7942 creates a high CO_2-requiring phenotype : evidence for a central role for carboxysomes in the CO_2 concentrating mechanism. *Plant Physiol* **91**, 505-513 (1989).

22. J. M. McGrath, S. P. Long, Can the cyanobacterial carbon-concentrating mechanism increase photosynthesis in crop species? A theoretical analysis. *Plant Physiology* **164**, 2247-2261 (2014).

23. A. Villarejo *et al.*, Evidence for a protein transported through the secretory pathway en route to the higher plant chloroplast. *Nature Cell Biology* **7**, 1224-1231 (2005).

24. C. V. Hoang, K. D. Chapman, Biochemical and molecular inhibition of plastidial carbonic anhydrase reduces the incorporation of acetate into lipids

in cotton embryos and tobacco cell suspensions and leaves. *Plant Physiology* **128**, 1417-1427 (2002).

25. N. U. Meldrum, F. J. W. Roughton, Carbonic anhydrase. Its preparation and properties. *The Journal of Physiology* **80**, 113-142 (1933).

26. B. S. Jacobson, F. Fong, R. L. Heath, Carbonic anhydrase of spinach: studies on its location, inhibition, and physiological function. *Plant Physiology* **55**, 468-474 (1975).

27. R. J. DiMario, M. C. Machingura, G. L. Waldrop, J. V. Moroney, The many types of carbonic anhydrases in photosynthetic organisms. *Plant Sci* **268**, 11-17 (2018).

28. A. T. Jagendorf, E. Uribe, ATP formation caused by acid-base transition of spinach chloroplasts. *Proceedings of the National Academy of Sciences* **55**, 170-177 (1966).

29. M. J. Mahon, pHluorin2: an enhanced, ratiometric, pH-sensitive green florescent protein. *Adv Biosci Biotechnol* **2**, 132-137 (2011).

30. C. T. Supuran, Structure and function of carbonic anhydrases. *Biochemical Journal* **473**, 2023-2032 (2016).

31. Q. Bruggeman, C. c. Raynaud, M. Benhamed, M. Delarue, To die or not to die? Lessons from lesion mimic mutants. *Frontiers in plant science* **6** (2015).

32. M. Manohar *et al.*, Identification of multiple salicylic acid-binding proteins using two high throughput screens. *Frontiers in plant science* **5** (2015).

33. R. A. Boyd, A. Gandin, A. B. Cousins, Temperature response of C4 photosynthesis: Biochemical analysis of Rubisco, Phosphoenolpyruvate Carboxylase and Carbonic Anhydrase in Setaria viridis. *Plant Physiology* 10.1104/pp.15.00586 (2015).

34. B. J. Nikolau, J. C. Hawke, Purification and characterization of maize leaf acetyl-coenzyme a carboxylase. *Archives of Biochemistry and Biophysics* **228**, 86-96 (1984).

35. D. Herbert *et al.*, Kinetic studies on two isoforms of acetyl-CoA carboxylase from maize leaves. *Biochemical Journal* **318**, 997-1006 (1996).

36. C. Alban, P. Baldet, R. Douce, Localization and characterization of two structurally different forms of acetyl-CoA carboxylase in young pea leaves, of which one is sensitive to aryloxyphenoxypropionate herbicides. *Biochemical Journal* **300**, 557-565 (1994).

37. V. Gueguen, D. Macherel, M. Jaquinod, R. Douce, J. Bourguignon, Fatty acid and lipoic acid biosynthesis in higher plant mitochondria. *Journal of Biological Chemistry* **275**, 5016-5025 (2000).

38. R. L. Lindroth, S.-Y. Hwang, "Diversity, redundancy, and multiplicity in chemical defense systems of aspen" in phytochemical diversity and redundancy in ecological interactions. (1996), 10.1007/978-1-4899-1754-6_2 chap. Chapter 2, pp. 25-56.

39. B. Winkel-Shirley, Biosynthesis of flavonoids and effects of stress. *Current Opinion in Plant Biology* **5**, 218-223 (2002).

40. F. J. Ferreira, C. Guo, J. R. Coleman, Reduction of plastid-localized carbonic anhydrase activity results in reduced arabidopsis seedling survivorship. *Plant Physiology* **147**, 585-594 (2008).

41. B. Gontero, J.-C. Meunier, J. Buc, J. Ricard, The 'slow' pH-induced conformational transition of chloroplast fructose 1,6-bisphosphatase and the control of the Calvin cycle. *European Journal of Biochemistry* **145**, 485-488 (1984).

42. G. D. Price, S. M. Howitt, Towards turbocharged photosynthesis. *Nature* **513**, 497 (2014).

43. B. M. Long *et al.*, Carboxysome encapsulation of the CO_2-fixing enzyme Rubisco in tobacco chloroplasts. *Nat Commun* **9**, 3570 (2018).

44. O. Emanuelsson, S. Brunak, G. von Heijne, H. Nielsen, Locating proteins in the cell using TargetP, SignalP and related tools. *Nat Protoc* **2**, 953-971 (2007).

45. I. A. Sparkes, J. Runions, A. Kearns, C. Hawes, Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. *Nat Protoc* **1**, 2019-2025 (2006).

46. M. Stemmer, T. Thumberger, M. Del Sol Keyer, J. Wittbrodt, J. L. Mateo, CCTop: an intuitive, flexible and reliable CRISPR/Cas9 target prediction tool. *PLoS One* **10**, e0124633 (2015).

47. C. Engler, R. Kandzia, S. Marillonnet, A one pot, one step, precision cloning method with high throughput capability. *PLoS One* **3**, e3647 (2008).

48. C. Engler, R. Gruetzner, R. Kandzia, S. Marillonnet, Golden gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes. *PLoS One* **4**, e5553 (2009).

49. E. Weber, R. Gruetzner, S. Werner, C. Engler, S. Marillonnet, Assembly of designer TAL effectors by Golden Gate cloning. *PLoS One* **6**, e19722 (2011).

50. C. S. Roberts, M. H. Spalding, Post-translational processing of the highly processed, secreted periplasmic carbonic anhydrase of Chlamydomonas is largely conserved in transgenic tobacco. *Plant Mol Biol* **29**, 303-315 (1995).

51. T. M. Stawicki *et al.*, The zebrafish merovingian mutant reveals a role for pH regulation in hair cell toxicity and function. *Dis Model Mech* **7**, 847-856 (2014).

52. A. L. Guo, G. Salih, D. F. Klessig, Activation of a diverse set of genes during the tobacco resistance response to TMV is independent of salicylic acid; induction of a subset is also ethylene independent. *Plant J* **21**, 409-418 (2000).

53. F. Les Erickson *et al.*, The helicase domain of the TMV replicase proteins induces the N-mediated defence response in tobacco. *The Plant Journal* **18**, 67-75 (1999).

54. H. S. Padgett, Y. Watanabe, R. N. Beachy, Identification of the TMV replicase sequence that activates the NGene-mediated hypersensitive response. *Molecular Plant-Microbe Interactions* **10**, 709-715 (1997).

55. H. S. Padgett, R. N. Beachy, Analysis of a tobacco mosaic virus strain capable of overcoming N gene-mediated resistance. *The Plant Cell* **5**, 577-586 (1993).

Primer	5'-3' Sequence	Purpose
SpeI dcTP_BCA1_For	CGCACTAGTATGGAAGAAATGGCTAACGAATCC	Remove cTP from βca1
BCA1_CA-DA For	CTGCCTCACGAGTGTGCCCATC	Q5 mutagenesis of zin binding sites
BCA1_CA-DA Rev	AGGCGGCAAAGACCATGTACTTG	Q5 mutagenesis of zin binding sites
BCA1_gRNA3_BsaIf	TGTGGTCTCAATTGTAGAGAACATTGTTGTCATGTTTTAGAGCTAGAAATAGCAAG	Create gRNA for βca1 off of pICH86966::AtU6p::sgRNA_PDS plasmid
BCA1_gRNA4_BsaIf	TGTGGTCTCAATTGTTGATTCAGAACCATCTGCGTTTTAGAGCTAGAAATAGCAAG	Create gRNA for βca1 off of pICH86966::AtU6p::sgRNA_PDS plasmid
BCA5_gRNA3_BsaIf	TGTGGTCTCAATTGCTAAAATTGAAAGCCTTGAGTTTTAGAGCTAGAAATAGCAAG	Create gRNA for βca5 off of pICH86966::AtU6p::sgRNA_PDS plasmid
BCA5_gRNA4_BsaIf	TGTGGTCTCAATTGACCCTAACCAAAGAAATGAGTTTTAGAGCTAGAAATAGCAAG	Create gRNA for βca5 off of pICH86966::AtU6p::sgRNA_PDS plasmid
pICH_bsa_FWD	AGGCACAGGTCTCGGGAGTGATCAAAAGTCCCACATCG	Adds RE sites to PCR product off of pICH86966
pICH_bsa_REV	AGGCACAGGTCTCGAGCGAAAAAAAGCACCGACTCG	Adds RE sites to PCR product off of pICH86967
pHlourin2-For	ATGGTGAGCAAGGGCGA	Amplifying pHlourin2 off pME plasmid
pHl2-Rev-EcoR1	ATTACGGAATTCGCCCTTTCACTTGTACAG	Amplifying pHlourin2 off pME plasmid. Adds EcoRI RE site
CTP-for-EcoR1	CATGGCGAATTCATGGATTCACAGCTAGTCTTGTCTCTG	Amplifying RecA transit peptide. Adds EcoRI RE site
CTP-pHl2-R1	TCGCCCTTGCTCACCATGTCGCGATCGAACTCAG	Amplifying RecA transit peptide. Aligns with pHlourin2-For and removes EcoRI site

Table 2.1. **Primers used in construct design**. Names and sequences of the primers used in the generation of plasmids for stable and transient transformation of tobacco.

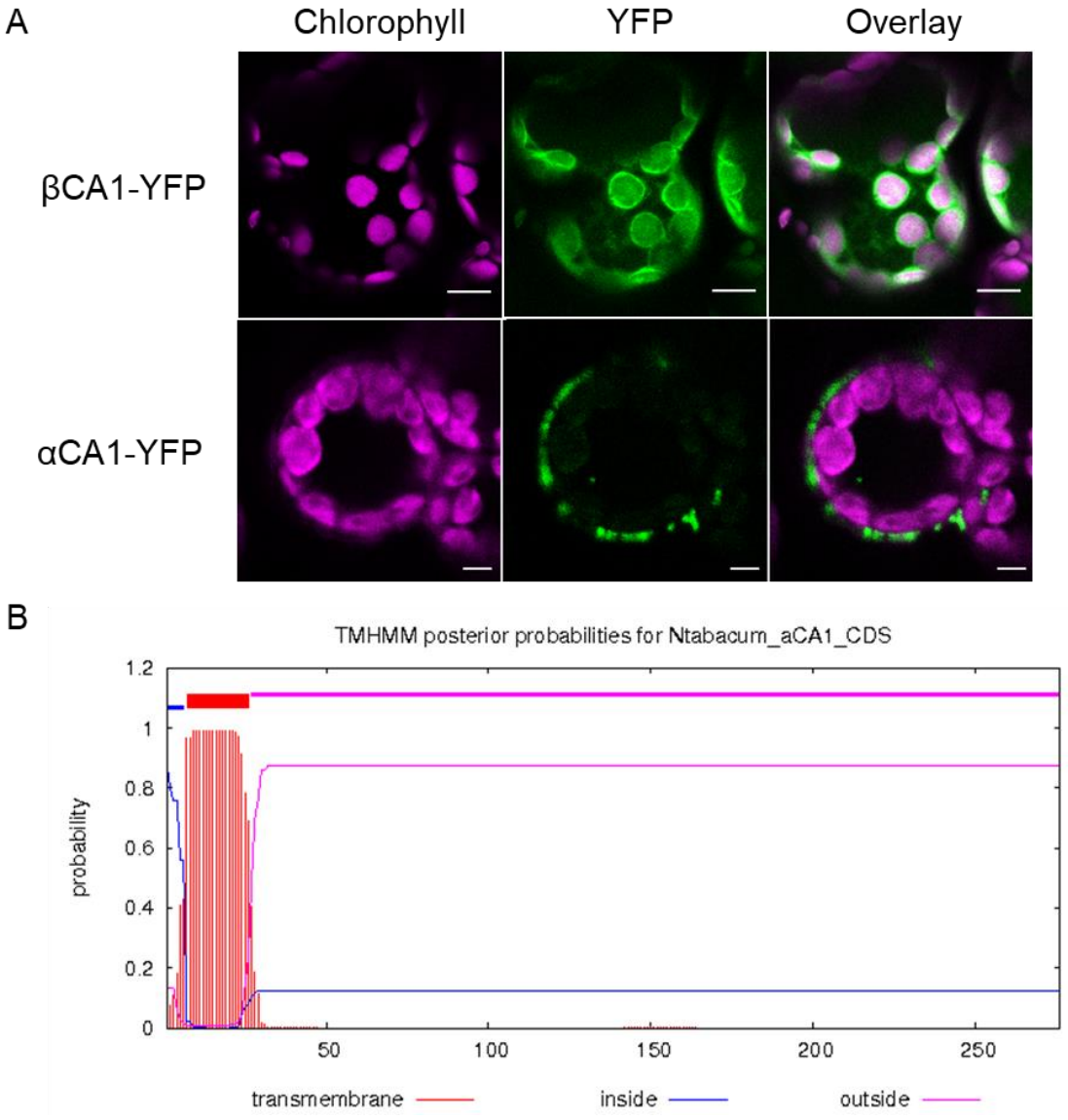

Figure S2.1. Localization of αCA1 in tobacco (A) Additional Confocal imaging of CA-YFP constructs transiently expressed in *N. tabacum* mesophyll cells. βCA1-YFP shows additional localization in the cytoplasm along with a "ring-like" formations around chloroplasts, likely indicating an import-bottleneck. αCA1-YFP is not found in the chloroplast stroma and appears to be located on the plasma membrane. Bars=5μm **(B)** TMHMM Server v. 2.0 prediction of αCA1 (accession XP_016483347) showing a transmembrane domain and peptide locations.

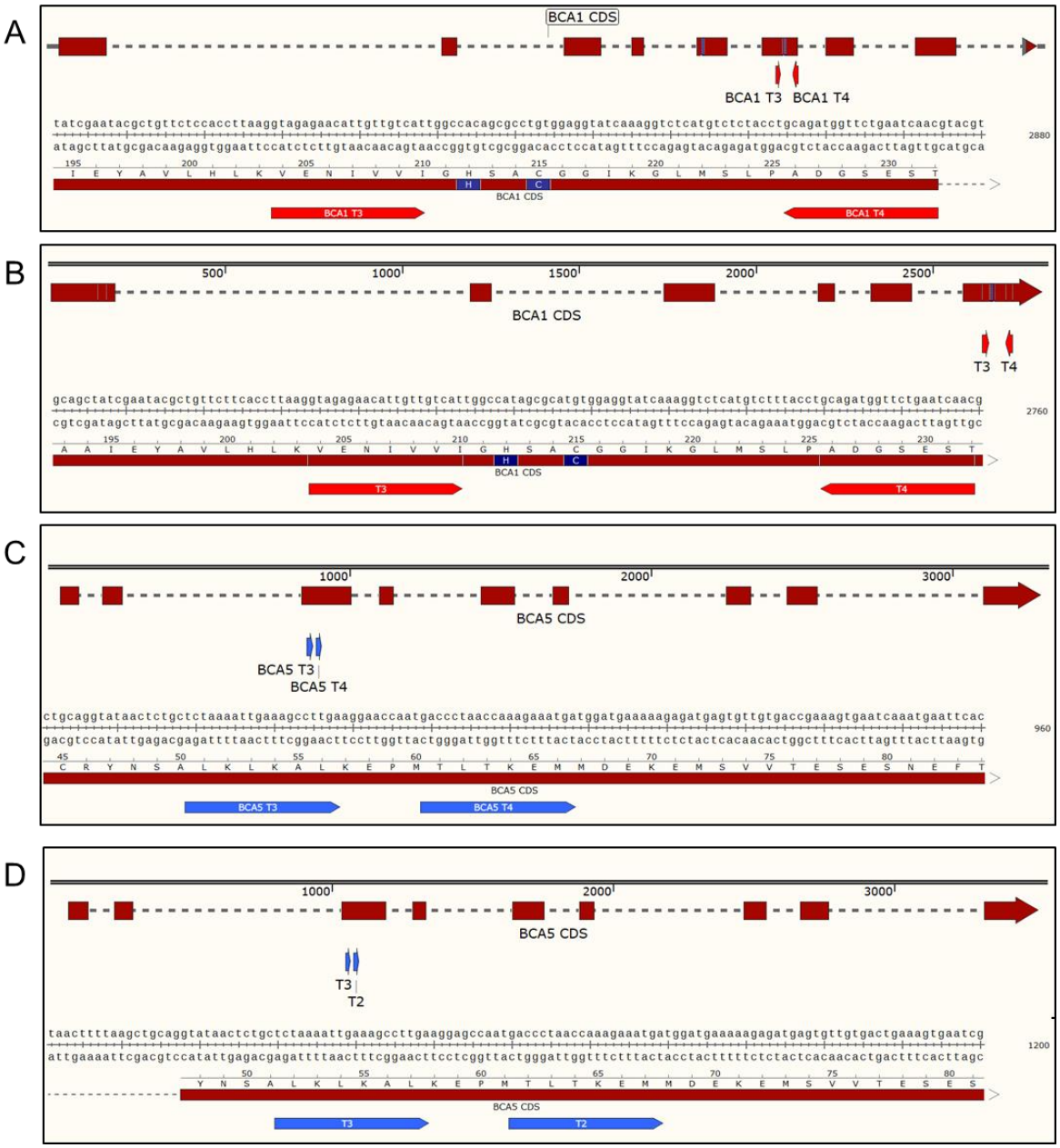

Figure S2.2 sgRNA targeting sites for βCA1 and βCA5. βCA1 CDS most similar to *N. sylvestris* (A) and *N. tomentosiformis* (B), which zinc-binding residues shown in blue. βCA5 CDS most similar to *N. sylvestris* (C) and *N. tomentosiformis* (D). Images made with SnapGene.

Figure S2.3. Sequencing of transgenic lines. Sequencing reaction results from T0 generation mutant lines created through the CRISPR/Cas9 system. Parenthesis indicate the net change in the nucleotide number between CA mutant homologs and the respective WT mRNA sequence. PAM sequences of each target are underlined. Sequences changes are indicated in red.

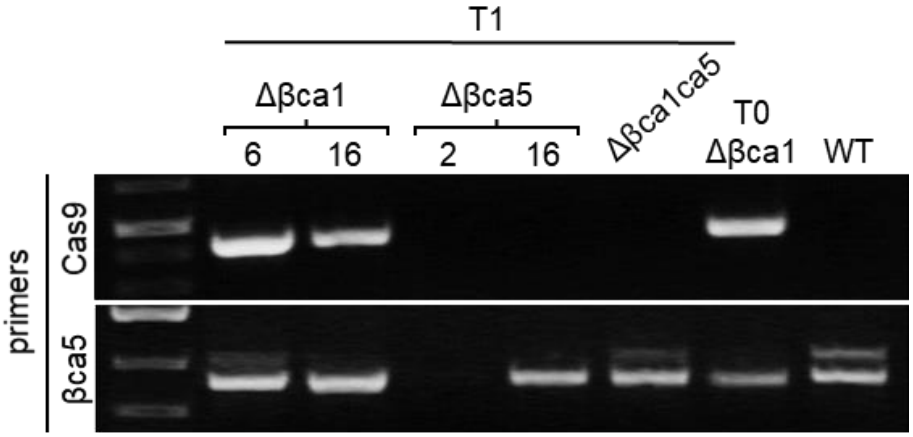

Figure S2.4. **Genotyping for Cas9⁻ mutants**. PCR (using gDNA template) of selected T1 plants (numbered) generated by self-pollination shows an absence of the CRISPR/Cas9 gene cassette in *Δβca1ca5*. This T1 *Cas9- Δβca1ca5* plant was used in the complementation experiments.

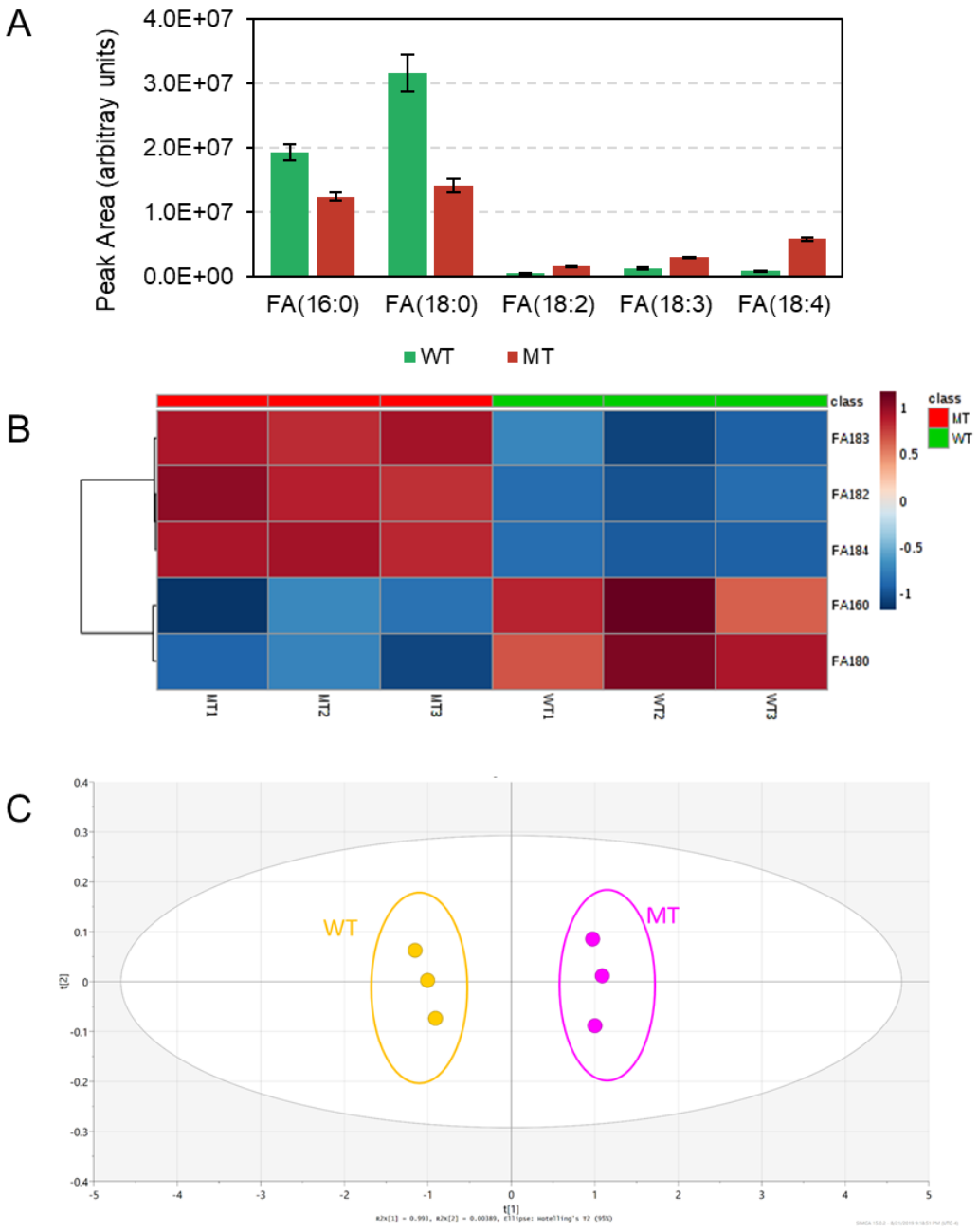

Figure S2.5. Free fatty acid analysis. (A) Peak area of WT and Δβ*ca1ca5* (MT) FFAs normalized with an internal standard (heptadecanoic acid) **(B)** Heat map of individual replicates using metaboanalyst data. **(C)** Principle component analysis score plot: Log transform and Pareto scale.

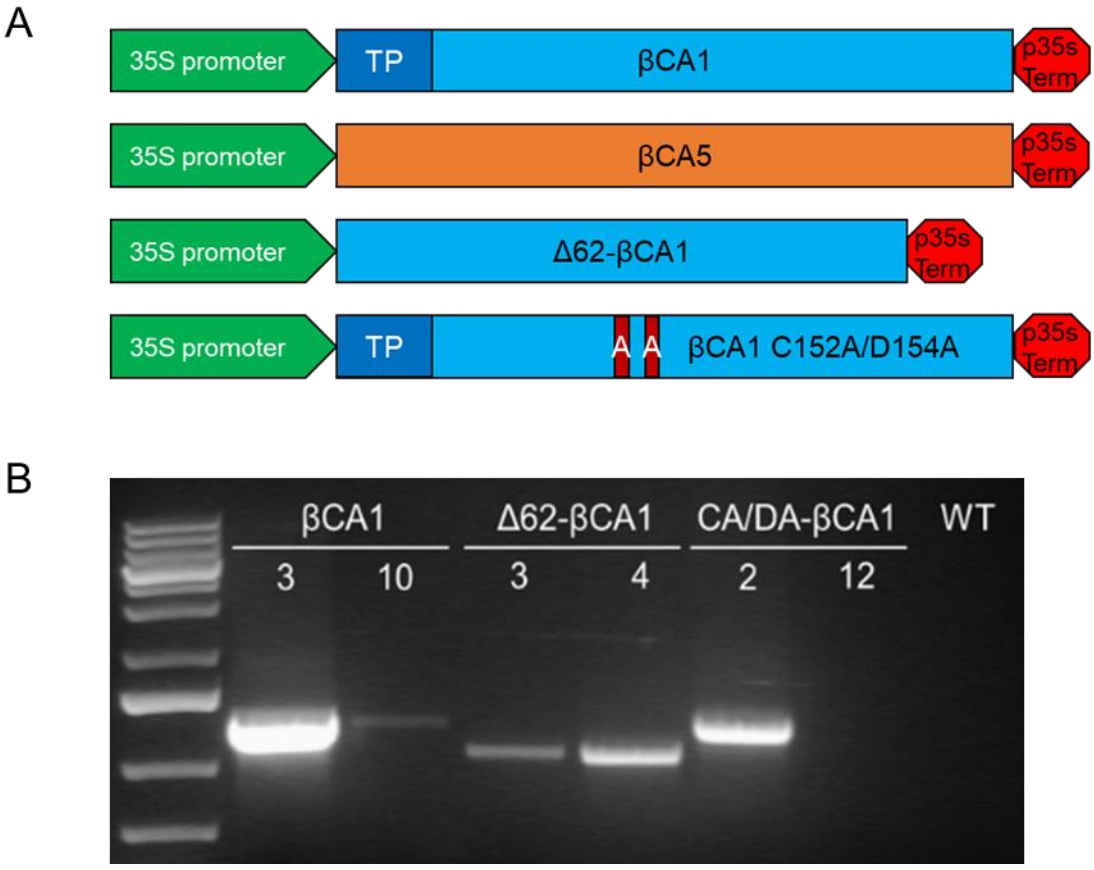

Figure S2.6. Complementation constructs and genotyping (A) Simplified schematic overview of complementation constructs**.** Each transgene was driving by a 35S overexpression promoter and used the 35S terminator. The chloroplast transit peptide for βCA5 is unknown. (B) Sample of genotyping data testing for the presence or absence of the plasmid insert.

Chapter 3

Bimolecular fluorescence complementation (BiFC) detects new interactions among β-carboxysomal proteins expressed in tobacco chloroplasts[6]

Abstract

The β-carboxysome is a proteinaceous bacterial microcompartment (BMC) that is an essential component of the carbon concentration mechanism of the cyanobacteria *Synechococcus elongatus* PCC7942. Current research aims to engineer β-carboxysomes into C_3 plants in order to improve photosynthesis by limiting the wasteful reaction ribulose 1,5-bisphosphate carboxylase/oxygenase has with oxygen. We used both traditional and multicolor bimolecular fluorescence complementation (BiFC) to study the interactions of β-carboxysomal proteins transiently expressed in the stroma of *Nicotiana benthamiana* chloroplasts. We were able to confirm the presence of previously reported protein-protein interactions of many β-carboxysomal proteins (CcmK2, CcmO, CcmL, and CcmN) as well as a new protein-protein

[6] Chapter presented as manuscript in preparation: Hines, KM., Lin, MT., and Hanson, MR. (2019) *Bimolecular fluorescence complementation (BiFC) detects new interactions among β-carboxysomal proteins expressed in tobacco chloroplasts.*

interaction previously unreported (CcmO with CcmN). We also observed that, while the last 17 C-terminal amino acids of CcmN are required to interact with CcmK2, they are not necessary for CcmN to interact with CcmO. Through multicolor BiFC, we observed that β-carboxysomal proteins CcmO and CccmL are able to form an interaction in the presence of a third β-carboxysomal protein, but not when they are expressed together in its absence. Our experiments establish a system to test the interactions between many carboxysomal proteins and provide insight into the interactions occurring within the structure or the carboxysome.

Introduction

Bacterial microcompartments (BMCs) are highly organized proteinaceous organelles found in a diverse range of bacterial species (1). The general structure of BMCs consists of an outer protein shell and internal structural proteins and enzymes (2). The protein oligomers that make up the shell contain pores may be selective for charged substrates (3, 4). A BMC provides multiple functions and advantages for the cells that house them. BMCs can increase the local concentration of substrates around enzymes and isolate toxic intermediates of certain enzymatic reactions (5, 6). One well-studied BMC is the β-carboxysome from the cyanobacteria *Synechococcus elongatus* PCC7942.

The β-carboxysome in *S. elongatus* (hereafter "carboxysome") contains much of the machinery for carbon fixation and functions as part of its carbon concentration mechanism (CCM) (2, 7, 8). The positively charged pores of the carboxysome shell allow negatively charged bicarbonate to enter the internal structure, where it interacts with the enzyme carbonic anhydrase and is converted to CO_2 and water (3, 9). The enzyme ribulose 1,5-bisphosphate carboxylase/oxygenase (RuBisCO) carries out the reaction of CO_2 and ribulose-1,5-bisphosphate (RuBP) inside the shell, creating 3-Phosphoglyceric Acid (PGA), which exits the carboxysome for the other downstream reactions (8).

The most up-to-date models of the carboxysome reveal to us that many specific interactions are required for formation of icosahedral bodies and proper function (7). CcmK2, CcmK3, and CcmK4 are single-domain BMC proteins (Pfam00936), which are able to oligomerize and form sheets of hexamers, according to crystal studies (4, 10). BMC proteins CcmO and CcmP contain tandem repeats of the single BMC domains and likely form into pseudo hexamer-like structures on the carboxysome facets (11). The facets on the carboxysome are formed by the interactions between the single and tandem BMC proteins (4, 10). At the vertices between the facets of the carboxysome shell is the CcmL pentamer (Pfam03319 protein domain), which has been shown to form an interaction with CcmK2 by FRET (12, 13). In our study, the proteins CcmK3, CcmK4, and CcmP were not investigated, as they appear to be redundant and are not absolutely required to form a functional carboxysome at sufficient CO_2 levels (14).

The internal proteins of the carboxysome include the structural proteins CcmM (with isoforms CcmM58 and CcmM35) and CcmN and enzymes CcaA (carbonic anhydrase) and the RuBisCO large and small subunits (7, 15). The larger CcmM isoform, CcmM58, has the ability to form direct interactions with the shell protein CcmK2, along with CcaA and CcmN (16-18). CcmN forms an interaction with CcmM58 and the CcmK2, the latter of which is mediated by its 18-amino acid encapsulation peptide at its C-terminus (18).

We have previously shown that transient expression of the major carboxysome shell proteins CcmK2, CcmO, and CcmL produces highly organized structures resembling empty carboxysomes in *Nicotiana benthamiana* chloroplasts (19). This observation is significant because the ability of the carboxysome proteins to form the correct interactions is pivotal for it to function properly as a CCM (20). In this report, we confirm, by the use of traditional and multicolor bimolecular fluorescence complementation (BiFC) that the correct interactions between the major carboxysomal shell proteins (CcmK2, CcmO, and CcmL) and the internal, structural proteins (CcmM58 and CcmN) are able to form within a new environment, *N. benthamiana* chloroplasts. BiFC allows us to test and visualize binding interactions of proteins in the novel environment of the chloroplast. We also used this BiFC system in the chloroplast environment to test previously undocumented protein-protein interactions among the carboxysomal proteins. Together these results shed light on the ability to engineer carboxysomes in chloroplast, provide new insights into the interactions that help form the carboxysome, and provide a new system to examine the interactions of carboxysome proteins and other proteins found in cyanobacteria.

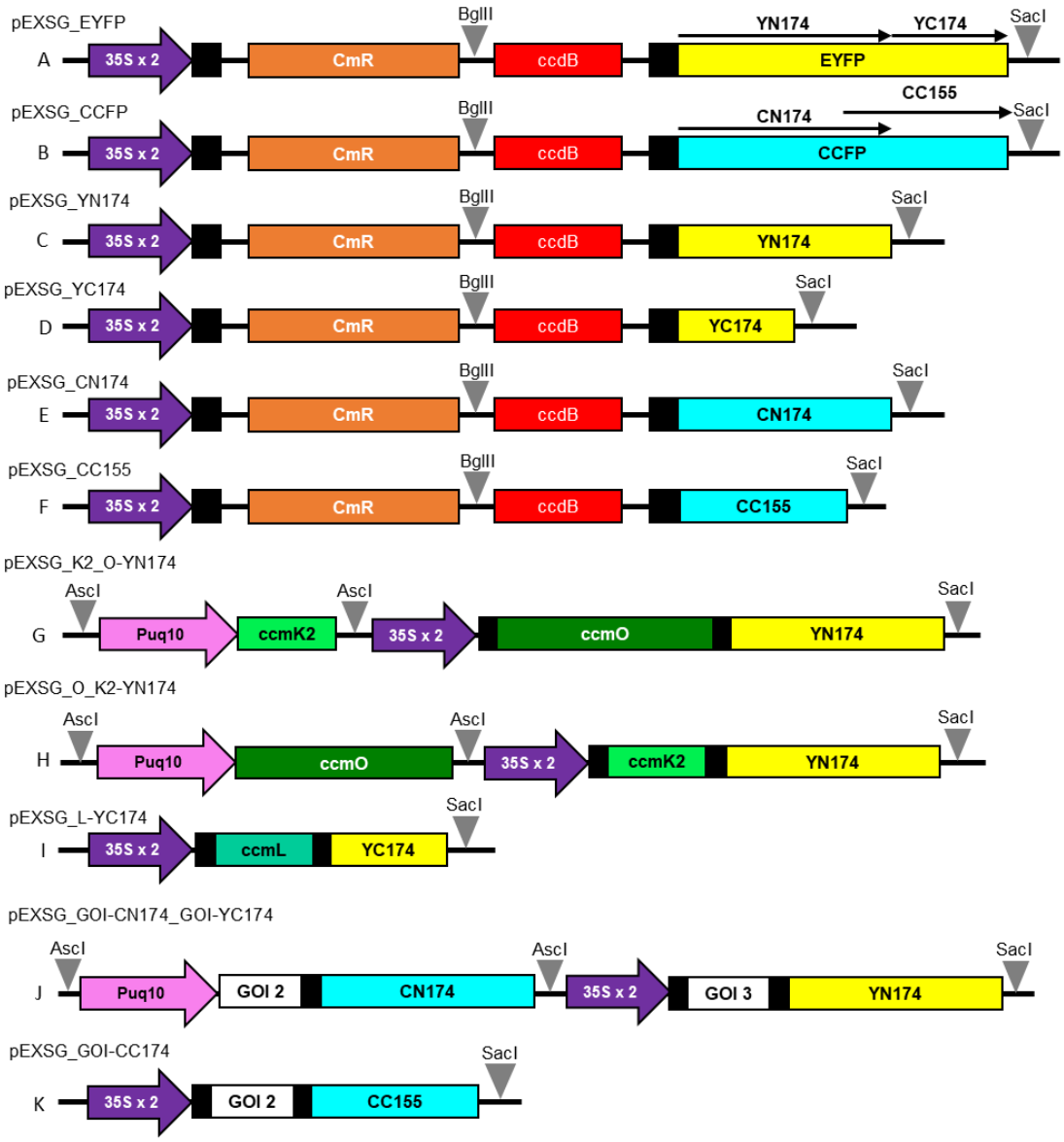

Figure 3.1. Construct design for BiFC plasmids. Shown: Region between the Left and right borders of the plant binary vector, pEXSG_EYFP and pEXSG_CCFP and their variants. Sections of the florescent proteins YFP and CCFP that were used to create the BiFC vectors are shown above with arrows (YN174, YC174, CN174, and CC155). The β-carboxysomal genes of interest are inserted from PCR8 vectors via an LR reaction between the two att recombination sites (Black boxes).The 35Sx2 Promoter consist of tandem repeats of the CaMV35s promoter (purple arrow) and the Puq10 promoter is the Arabidopsis ubiquitin 10 promoter (pink arrows). The GOI (white) regions represent the different combinations included into the multicolor BiFC vectors. Important restrictions sites used in construction of the BiFC plasmids are marked with triangles.

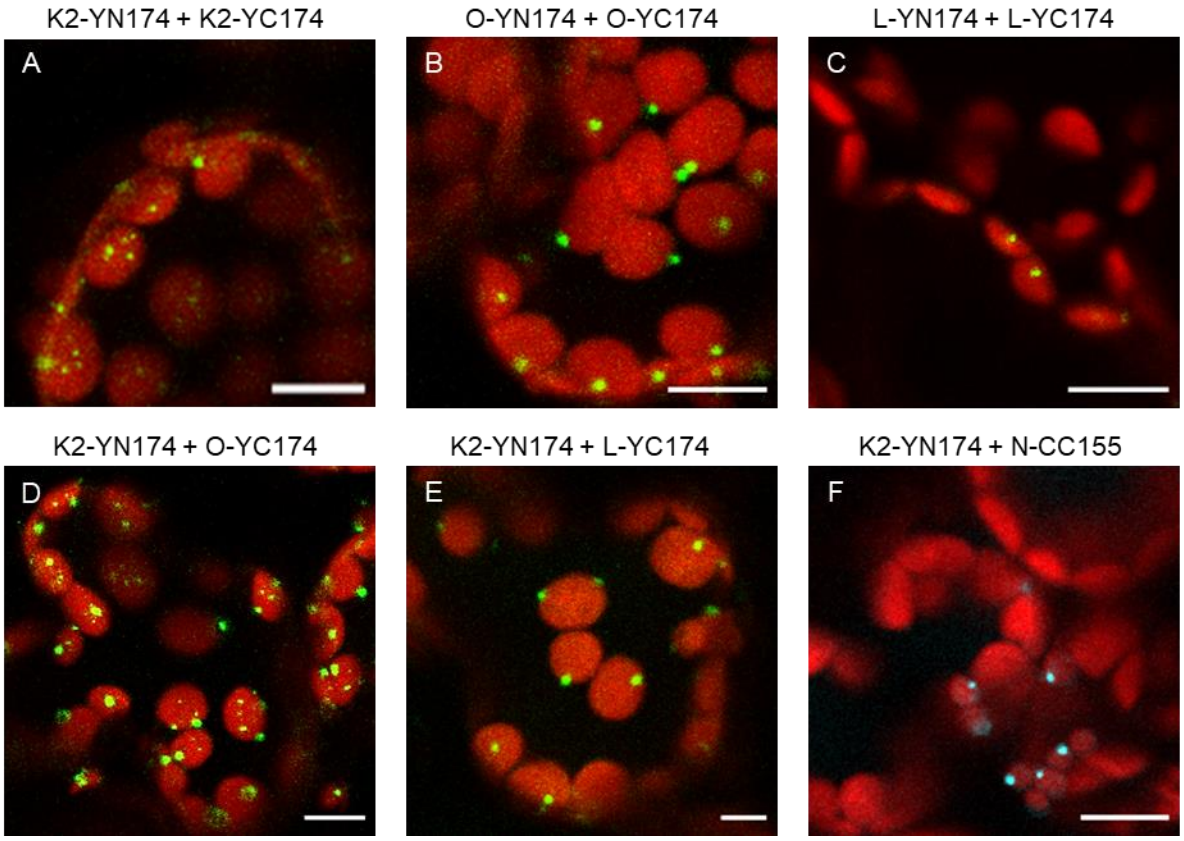

Figure 3.2. BiFC confocal images of β-carboxysomal proteins with known interactions. **A:** pEXSG_K2-YN coexpressed with pEXSG_K2-YC. **B:** pEXSG_O-YN with pEXSG_O-YC. **C:** pEXSG_L-YN with pEXSG-L_YC. **D:** pEXSG_K2-YN with pEXSG_O-YC. **E:** pEXSG_K2-YN with pEXSG_L-YC. **F:** pEXSG_K2_YN with pEXSG_N-CC. Green: BiFC YFP signal. Cyan: BiFC CFP signal. Red: Chlorophyll A. Bars 10µm

Results

The β-carboxysomal proteins described below have been previously fused with the Arabidopsis *recA* chloroplast transit peptide at their N-terminal ends (19, 21) for targeting to the chloroplast stroma.

Split YFP and CFP results of known interactors

The interactions between individual carboxysome proteins have been studied in different systems and have led to detailed models of the carboxysome structure (7). In this first section, we examine protein-protein interactions that have been observed in other systems.

One of the three major shell proteins, CcmK2, (hereafter K2), has been shown to self-polymerize (22) and form into hexamers in the facets of the carboxysome shell (4). We used BiFC with split Enhanced yellow fluorescent protein (EYFP) to test the specific interactions occurring between the K2 subunits in the chloroplasts of *N. benthamiana*. K2 was inserted into two modified plant binary vectors pEXSG_YN174 and pEXSG_YC174 (Figure 3.1, between the attR regions) and transiently expressed in *N. benthamiana* leaves. The confocal images show bright punctate loci in the chloroplasts indicating a K2-K2 interaction (Figure 3.2). CcmO (hereafter O) is a tandem-BMC shell protein believed to form a trimer to and resemble a pseudo-hexameric unit (11). Following the same procedures as with K2, split YFP BiFC shows interaction between the O subunits in the chloroplast (Figure 3.2B). CcmL (hereafter L) is

a shell protein making up the vertices of the carboxysome and is able to form pentameric structures (12, 13). Our split YFP results show that the L subunits are able to self-polymerize within the chloroplast (Figure 3.3).

Other known interactions among the shell proteins include K2's interaction with the other two shell proteins O and L (13, 14). Both of these interactions were observed in chloroplast in the form of small punctate loci with K2 and O (Figure 3.2D) and large polar aggregates when K2 and L are expressed (Figure 3.2E). A small 18-amino acid C-terminal sequence on CcmN (hereafter N) termed the encapsulation peptide mediates its association with the shell protein K2 (18, 19). Here, using pEXSG plasmids with split cerulean cyan fluorescent protein (CCFP) (Figure 3.1E, F) we observed interactions between K2 and N when co-expressed (Figure 3.2F).

Undocumented Interactions

Our system provides unique opportunities to study carboxysomal proteins. Due to the evolutionary history of plastids, the environment of the chloroplast is more similar to cyanobacteria than yeast but may lack possible interacting proteins or chaperones that might be present in the cytoplasm of cyanobacteria. This allows us to test previously undocumented interactions of carboxysomal proteins in the absence of endogenous cyanobacterial proteins.

The encapsulation peptide of N has been shown to be necessary for its association with K2, but it is not known whether N interacts with O or, if it

does, whether this 18 amino acid sequence is required. Using split CCFP BiFC pEXSG plasmids (Figure 3.1E, F), O and N were co-expressed in *N. benthamiana*. We observed many puncta within the chloroplast showing a clear interaction between O and N (Figure 3.3A). We next tested to see if the encapsulation peptide required for the N-K2 interaction is also required for N's observed interaction with O. We used a mutant form of N in which we removed the last 18 codons (termed NΔ18) to co-express with O. When O and NΔ18 were co-expressed using split-YFP BiFC, we observed a diffuse YFP signal in the chloroplasts. The excitation of chloroplasts naturally creates a background YFP-like signal, so we measured the fluorescent intensity of chloroplasts that appeared bright yellow and chloroplasts appearing mostly red (Figure 3.3E, F). The fluorescent intensity of the YFP signal in chloroplast of transformed cells was much greater than the background YFP signal observed in untransformed cells. K2 was co-expressed with NΔ18 as a negative control (Figure 3.3C) and no interaction was observed. Some proposed models show O interacting with L at the vertices of the carboxysome structure (14), but co-expression of O and L split-YFP plasmids yielded no positive signal (Figure 3.3D).

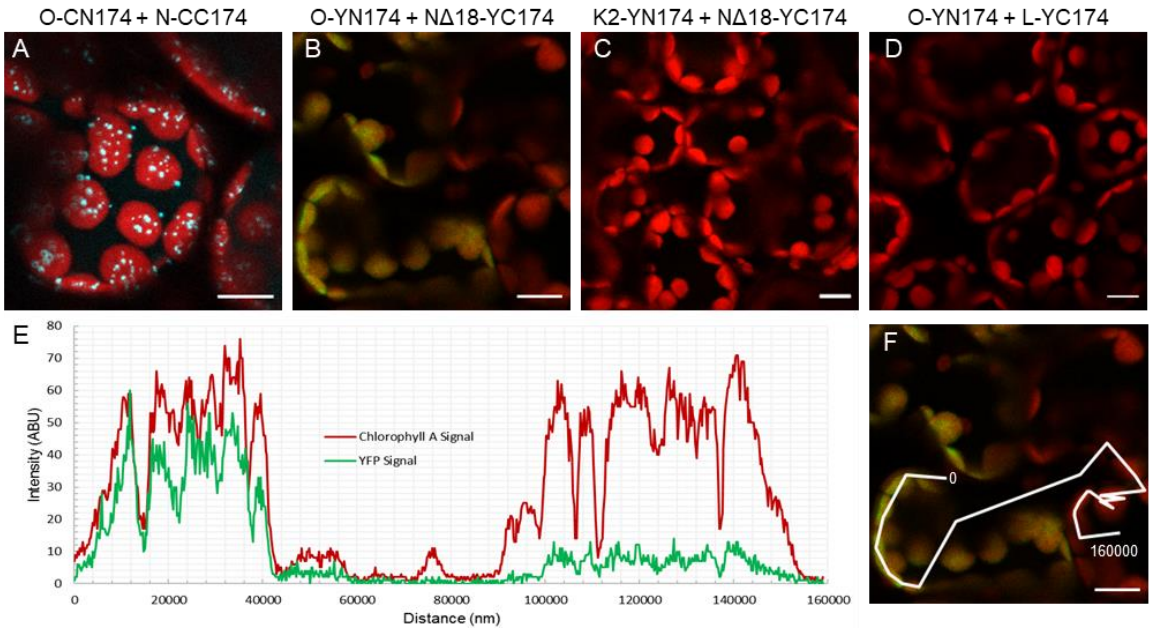

Figure 3.3. BiFC confocal images of β-carboxysomal proteins testing undocumented interactions. A: pEXSG_O-CN coexpressed with pEXSG_N-CC. **B:** pEXSG_O-YN with pEXSG_NΔ18-YC. **C:** pEXSG_K2-YN with pEXSG-NΔ18_YC. **D:** pEXSG_O-YN with pEXSG_O-LC. **E:** Graph comparing the fluorescence intensities between the diffused YFP signal in B with the background signal in Arbitrary Brightness Units (ABU) Laser intensities: 633nm: 0.96%, 514nm: 0.25%. **F**: Trace of measurements in pEXSG_O-YN with pEXSG_NΔ18-YC coexpression. Green: BiFC YFP signal. Cyan: BiFC CFP signal. Red: Chlorophyll A. Bars 10µm

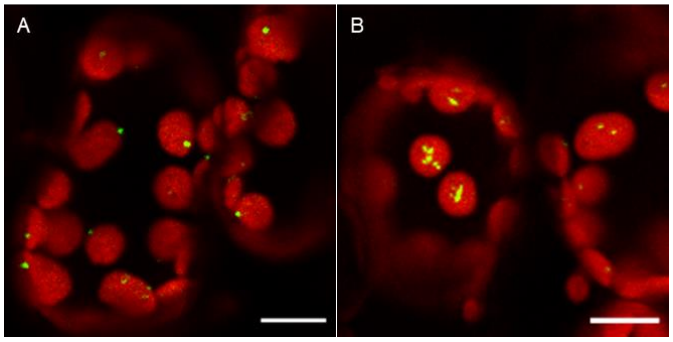

Figure 3.4. BiFC confocal images of the three major β-carboxysomal shell proteins coexpressed in *N. benthamiana* chloroplasts; **A.** pEXSG_O_K2-YN coexpressed with pEXSG_L-YC. **B.** pEXSG_K2_O-YN coexpressed with pEXSG_L-YC. Green: BiFC YFP signal. Red: Chlorophyll A. Bars 10µm.

Triple Expression

Due to the sequence similarity of O and K2 (23), the lack of interaction between O and L was surprising. To further explore the nature of the interactions between the shell proteins, a BiFC system which expresses three proteins (one of which is unlabeled) was created (Figure 3.1G-I). Unlabeled K2 and O, driven by the *Arabidopsis* ubiquitin 10 promoter, was added to the AscI site upstream of O-YN174 and K2-YN174, respectively, on the pEXSG vector. These constructs, when co-expressed with L-YC174, showed that L interacts with either O or K2 when the three shell proteins are all present. Confocal images indicate that L is able to interact with K2 in the presence of O as it did in the absence of O (compare Figure 3.4 and Figure 3.2E). In both cases, bright aggregates were observed. While no interaction between O and L was seen when the two proteins were expressed without K2 (Figure 3.3D), multiple bright puncta were observed when O and L were co-expressed together in the presence of K2 (Figure 3.4B).

To expand our understanding of these interactions further, a multicolor BiFC system was created using the pEXSG vectors that labels two proteins with larger CN174 and YN174 split proteins along with a third protein labeled with CC155 (Figure 3.1J and 3.1K), which is able to create a fluorescent signal with both of the larger split proteins (24). An interaction between two proteins that brings the split YN174 and CC155 together produces an YFP-like signal, and

an interaction between CN174 and CC155, produces a CFP-like signal. This system allows us to observe directly any interactions between three proteins.

The interaction between the three shell proteins was tested with the new system. When K2 was co-expressed with L and O, K2 showed the ability to interact with both proteins simultaneously (K2-L interaction in CFP channel and K2-O in YFP channel), and these interactions appeared to be co-localized (Figure 3.5). O was able to form an interaction with both L and K2 when the three proteins were co-expressed (Figure 3.5B), again forming co-localized puncta. Lastly, L was able to form interactions with K2 and O when co-expressed, forming the same co-localized puncta in the chloroplast that the other multicolor BiFC constructs created (Figure 3.5C). These confocal data appear to show that all three major shell proteins are able to form interactions with one another when they are expressed and targeted to the chloroplast, but the possibility remains that the BiFC signal between O and L could be due to proximity of the O and L in the presence of K2 rather than the direct interaction between O and L.

We next expanded this multi-color BiFC system to test the interactions between the shell proteins and two internal structural proteins N and CcmM58 (hereafter M58). Plasmids were created to co-express N-CN and M58-YN with a third protein containing the C-terminal CFP peptide. We also used the

previously designed pEXSG plasmids used to create the expression patterns in Figure 3.5.

The shell protein K2 was able to interact with both N and M58 in the chloroplast when all three were co-expressed (Figure 3.6); however, the interaction between K2-N (the CFP channel) appears to produce more puncta than the M58-N interaction. O was also able to interact with both N and M58 (Figure 3.6B) and, like K2, appears to favor an interaction with N over M58.

N showed strong interactions with both L and K2 when all three were co-expressed (Figure 3.6C), and when with both K2 and O were expressed (Figure 3.6D). In both cases, bright punctate loci were observed in both YFP and CFP channels. When M58 was co-expressed with both L and K2, very dim YFP (diffuse) and CFP signals were observed. The interaction of M58 with the shell proteins is not necessarily surprising; while N appears to be required for encapsulation of the internal enzymes and proteins of the carboxysome, the co-expression of M58 with K2 and O was able to produced highly organized structures resembling empty microcompartments in *N. benthamiana* (19).

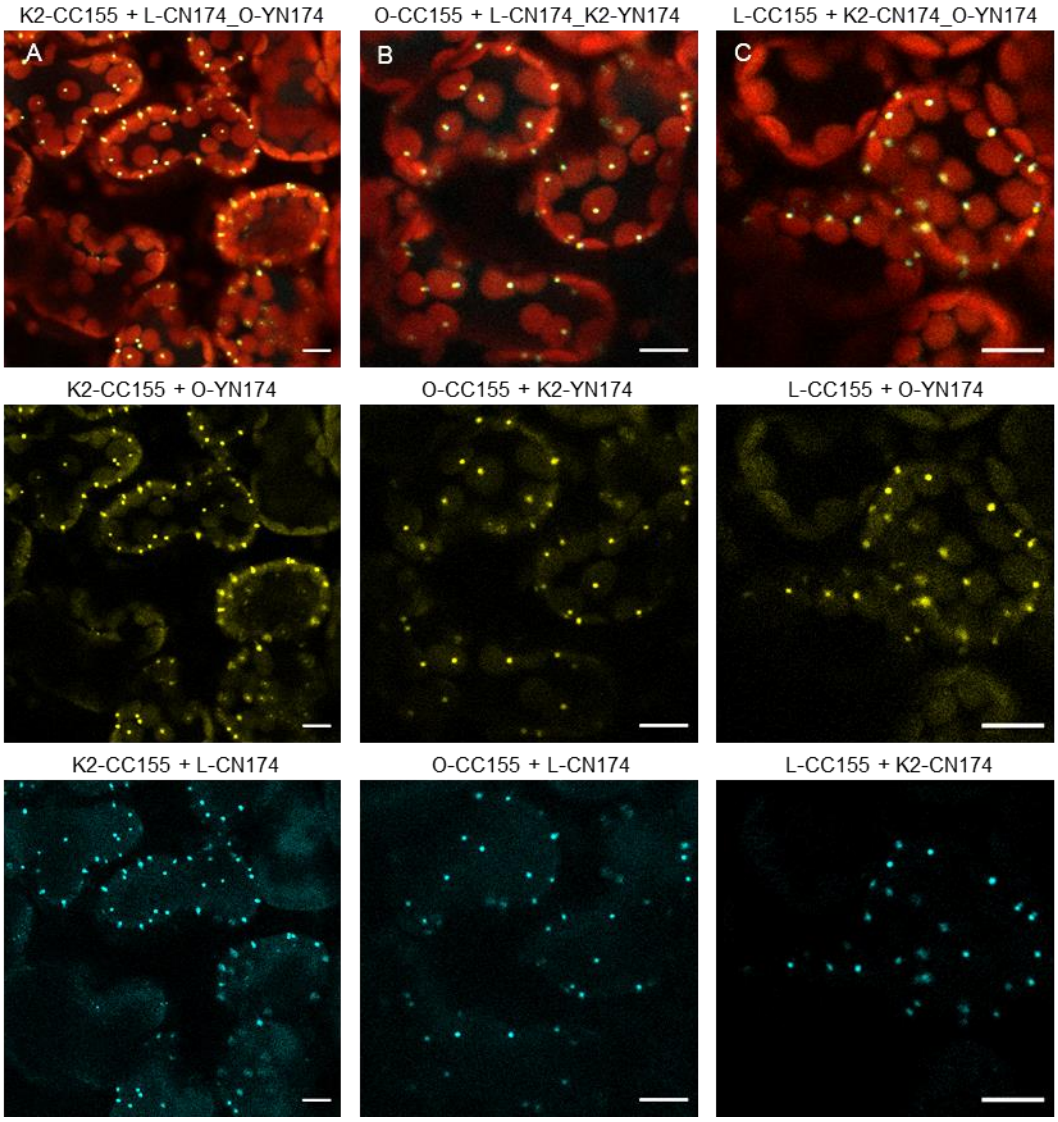

Figure 3.5. Confocal images of multicolor BiFC between the carboxysome shell proteins. **A.** pEXSG_K2-CC155 coexpressed with pEXSG_L-CN174_O-YN174 **B.** pEXSG_O-CC155 with pEXSG_L-CN174_K2-YN174. **C.** pEXSG_L-CC155 with pEXSG_K2-CN174_O-YN174. Titles above images indicate the two proteins shown to be interacting in that channel. Top row: Merged (including Chlorophyll A channel). Middle row: YFP signal. Bottom row: CFP signal. Yellow: BiFC YFP signal. Cyan: BiFC CFP signal. Red: Chlorophyll A. Bars 10μm

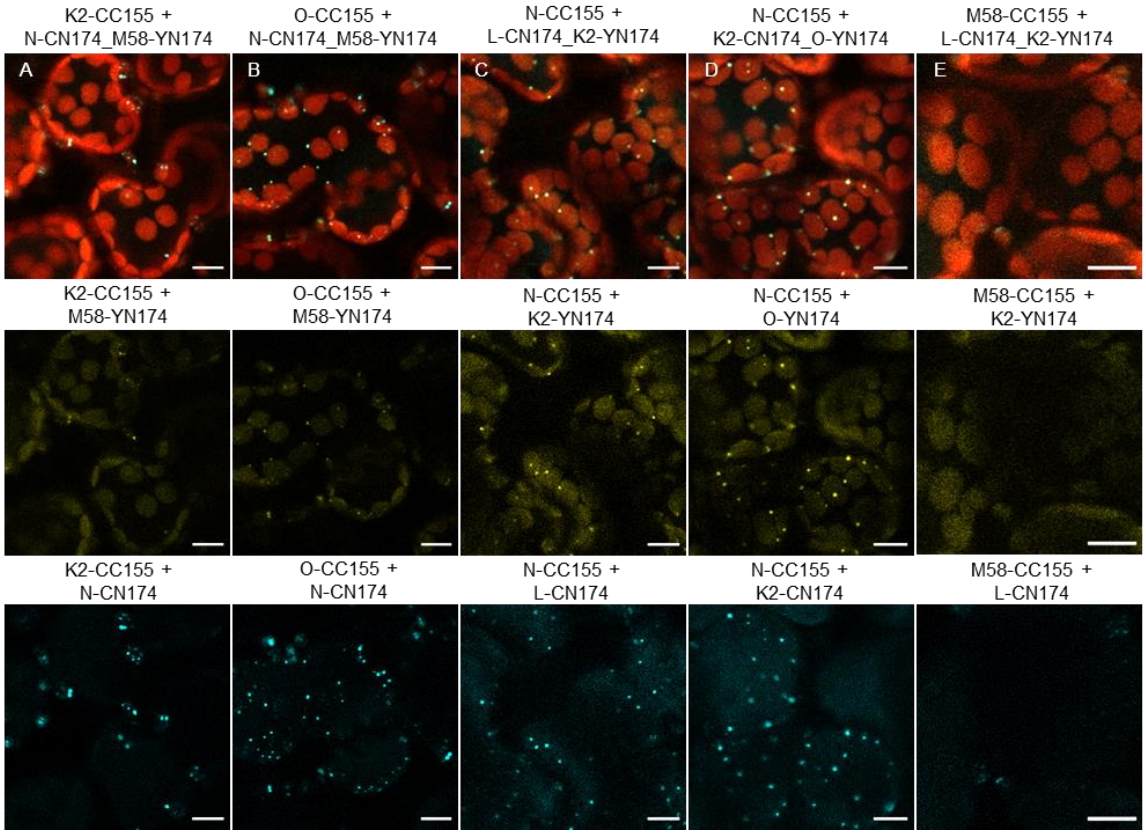

Figure 3.6. Confocal results of multicolor BIFC results between the shell and internal, structural proteins of the β-carboxysome. Titles above images indicate the two proteins shown to be interacting in that channel. Top to bottom: Merged (including Chlorophyll A channel), YFP signal, and CFP signal. **A.** Plasmid pEXSG_K2-CC155 coexpressed with pEXSG_N-CN174_M58-YN174, **B.** pEXSG_O-CC155 coexpressed with pEXSG_N-CN174_M58-YN174, **C.** pEXSG_N-CC155 with pEXSG_L-CN174_K2-YN174, **D.** pEXSG_N_CC155 with K2-CN174_O_YN174, **E.** pEXSG_M58-CC155 with pEXSG_L-CN174_K2-YN174. Yellow: BiFC YFP signal. Cyan: BiFC CFP signal. Red: Chlorophyll A. Bars 10μm

Discussion

Introducing a CCM into chloroplasts is an attractive solution to the inefficiencies of C_3 photosynthesizing plants (25); however the apparent step-wise assembly process of β-carboxysomes has been a cause for concern for the goal of engineering them to the chloroplast (20). While α-carboxysomes have been assembled in *E. coli* (26), there has been no report of expression of β-carboxysomal proteins outside of their native cyanobacteria to form a functional microcompartment. Recent studies have provided insight into the possibility of assembling a β-carboxysome into chloroplast, as the three major shell proteins self-assembled into an organized structure inside the chloroplast of tobacco (19).

Previous reports on BMC protein-protein interactions largely were performed in yeast two-hybrid systems or *in vitro*. Testing the interactions between carboxysomal proteins in chloroplasts provides an opportunity to observe them in a pH environment that is highly similar to cyanobacteria (27, 28), but free of other potential interactors. Combining BiFC with the transient transformation of tobacco allows testing previously unknown protein-protein interactions of the carboxysome which, in turn, can lead to better model determination of the overall structure.

Testing protein-protein interactions with BiFC is not without caveats. The ability of co-expressed proteins to form multiple punctate loci was

sometimes observed in this study (example in Figure 3.3A), but expression of certain constructs resulted in the form of a single punctum. It is possible that these aggregates could result from non-specific binding caused by the misfolding of certain carboxysomal proteins and may not be indicative of real protein-protein interactions that occur *in vivo*. When proteins misfold, they tend to have exposed nonpolar side chains, which can lead to a large aggregation of the misfolded proteins (29, 30). With these proteins in close proximity to one-another, the split fluorescent proteins could be close enough to come together to give a positive looking signal. Due to the nature of BiFC, interactions are permanently held together by the strong binding that occurs between the beta-sheets of the split CFP and YFP proteins (31), therefore only a brief encounter in protein aggregates may be sufficient to observe a positive signal.

Primer Name	5'-3' Sequence
pEX-Yfor	GCGTAAAGATCTGGATCCGGCTTA
pEX-Ynrev	CTCTAGAGCTCTTAGTCCTCGATGTTGTGGCGGAT
pEX-YC1rev	GCGAGCTGCACGCTGCCCCCATCAACCACTTTGTACAAGAAAG
pEX-YC2for	GGCAGCGTGCAGCTCGC
pEX-YC2rev	GACTCTAGAGCTCTTACTTGTACAG
pEX-CC1rev	ATGCCGTTCTTCTGCTTGTCCCCATCAACCACTTTGTACAAGAAAG
pEX-CC2for	GACAAGCAGAAGAACGGCAT
pEX4556rev	GCTCAACACATGAGCGAAACC

Table 3.1. Primers used to create the BiFC and multicolor BiFC plasmids. The names and specific 5' to 3' sequences as shown.

Materials and Methods

Plant expression BiFC vector design

BiFC constructs were designed based on Hu and Kerppola (24). To create pEXSG_YN174 (Figure 3.1), the plant binary vector pEXSG_EYFP was amplified by PCR with primers pEX-Yfor and pEX-Ynrev (Table 3.1). The PCR product and pEXSG_EYFP were digested with Fast Digest enzymes BglII and SacI (Life Technologies Carlsbad, CA), run on an agarose gel and purified (the larger band of the pEXSG_EYFP digest was selected). Digest products were ligated with T4 DNA ligase (Life Technologies) and transformed into One Shot Top10 cells (Life Technologies) and plated on LB plates with ampicillin.

To create pEXSG_YC174 (Figure 3.1), pEXSG_EYFP was amplified by PCR in a three step, overlapping protocol. Primers pEX-Yfor and pEX-YC1rev (Table 3.1) were used in the 1st round of PCR to amplify the BglII to *ccdB* region with overhangs matching the pEX-YC2for primer and the start of the C-terminal EYFP protein (YC174). The 2nd round of PCR used primers pEX-YC2for and pEX-YC2rev (Table 3.1) to amplify the YC174 portion of EYFP to the SacI site on pEXSG_EYFP. The PCR products from both reactions were then used as the template in a 3rd round of PCR with primers pEX-Yfor and pEX_YC2rev to create a PCR product that contained RE sites BglII and SacI along with *ccdB* and YC174. This was then digested with enzymes BglII and SacI and ligated into pEXSG_EYFP previously digested with the same

enzymes and then transformed into Top10 cells and plated on LB plates containing ampicillin.

The multi-expression BiFC plasmids (Figures 3.4, 3.5, and 3.6) were generated by digesting the plasmids published in Lin, *et al.* (19), which contained unlabeled K2, O, and L driven by the Arabidopsis ubiquitin 10 promoter (Puq10) that was flanked by two AscI sites. Those plasmids along with pEXSG_K2-YN174¸ pEXSG_O-YN174, and pEXSG_L-YN174 were digested with AscI (Life Technologies Cat. Num. FD1864) and ligated with T4 DNA ligase to create the plasmids seen expressed in Figure 3.4.

To create multicolor BiFC plasmids pEXSG_CN174 and pEXSG_CC155 were created from pEXSG_CCFP based on Hu and Kerppola (24). CC155 is able to interact with CC174 and YN174 giving either a CFP-like signal or an YFP-like signal. Due to the similarity in the sequences, the plasmid pEXSG_CN174 was created using the same primers used to make pEXSG_YN174. As before, the C-terminus-containing plasmid was created using three rounds of PCR. The primers pEX-Yfor and pEX-CC1rev (Table 3.1) were used PCR to amplify the BglII site and the ccdB region of the pEXSG_CCFP plasmid and the primers pEX-CC2for and pEX-YC2rev were used to amplify the C-terminal portion of CCFP and the Saci site on the plasmid. These two PCR products were then used as templates for the 3rd round of PCR and joined in a reaction using pEX-Yfor and pEXYC2rev. The PCR

products for CC174 and CC155 were digested and ligated as previously describe for pEXSG_YN174 and pEXSG_YC174.

To create the multicolor BiFC plasmids, a stepwise overlapping PCR protocol was used again. A pEXSG_GOI-CN174 plasmid (where GOI is a specific β-carboxysomal protein) was amplified by PCR with PU10CTPf, which aligns with the RecA chloroplast transit peptide and contains an overhang matching the Puq10 promoter, and pA35Sr (adds an AscI site in its overhanging region). The promoter Puq10 was amplified with PCR off a different plasmid with UQ10P2re5 and Pub10r which adds an AscI site and aligns with PU10CTPf respectively. The two PCR products were used as templates for an overlapping, fusion PCR with UQ10P2re5 and pA35Sr. The final PCR product and a pEXSG_GOI-TN174 plasmid were digested with AscI, purified, and then ligated.

Transient transformation of N. benthamiana leaves

Liquid cultures of *Agrobacterium tumefaciens* GV3101RK transformed with pEXSG plasmids were centrifuged, re-suspended in 1.5 mL of 10mM $MgCl_2$, 10mM MES, PH5.6 solution, and then centrifuged again. The cells were then re-suspended with 10mM $MgCl_2$, 10mM MES, pH5.6 with 1M acetosyringone (5000x) and the OD_{600} was measured. Cells were diluted to OD_{600} 1.0 and kept in the dark for 3 hours at 28C for 3-4 hours (32). To co-express two plasmids, a 1:1 ratio of the two cell lines were mixed (equating to

an OD_{600} of 0.5 for each transformed cell line) by pipetting and infiltrated into the leaves of *N. benthamiana* together at 0.8mL per leaf. The plants were kept on a short-day light cycle (12 hours light, 12 hours dark) for 2 days and then imaged.

Confocal microscopy

Leaf samples were imaged on the Zeiss LSM 710 AxioObserver confocal microscope using a 25X water emersion objective (LD LCI Plan-Apochromat 25x/0.8 Imm Korr DIC M27). As described in Hu and Kerppola (24), the excitation and emission spectra of multicolor BiFC YFP and CFP signals are shifted. The emission created from YN174-CC155 is more like GFP than YFP. The 405nm laser was used to excite CN174-CC155 and the 488nm laser was used for YN174-CC155. For the BiFC experiments involving only the split YFP, the 514nm laser was used because the emission and excitation spectra are more similar to YFP.

Acknowledgements

We thank Dr. Cheryl Kerfeld (Michigan State) for providing us with the vectors harboring the *ccmK2*, *ccmL*, and *ccmO* genes, and the genomic DNA extracted from *Synechococcus elongatus* PCC7942. We also acknowledge support from the Cornell University Biotechnology Resource Center (National Institutes of Health [grant number S10RR025502]) for the shared Zeiss LSM 710 confocal microscope.

References

1. T. O. Yeates, Y. Tsai, S. Tanaka, M. R. Sawaya, C. A. Kerfeld, Self-assembly in the carboxysome: a viral capsid-like protein shell in bacterial cells. *Biochem Soc Trans* **35**, 508-511 (2007).
2. T. O. Yeates, C. A. Kerfeld, S. Heinhorst, G. C. Cannon, J. M. Shively, Protein-based organelles in bacteria: carboxysomes and related microcompartments. *Nat Rev Microbiol* **6**, 681-691 (2008).
3. Z. Dou *et al.*, CO2 fixation kinetics of *Halothiobacillus neapolitanus* mutant carboxysomes lacking carbonic anhydrase suggest the shell acts as a diffusional barrier for CO_2. *J Biol Chem* **283**, 10377-10384 (2008).
4. C. A. Kerfeld *et al.*, Protein structures forming the shell of primitive bacterial organelles. *Science* **309**, 936-938 (2005).
5. C. A. Kerfeld, S. Heinhorst, G. C. Cannon, Bacterial microcompartments. *Annu Rev Microbiol* **64**, 391-408 (2010).
6. J. T. Penrod, J. R. Roth, Conserving a volatile metabolite: a role for carboxysome-like organelles in *Salmonella enterica*. *J Bacteriol* **188**, 2865-2874 (2006).
7. B. D. Rae, B. M. Long, M. R. Badger, G. D. Price, Functions, compositions, and evolution of the two types of carboxysomes: polyhedral microcompartments that facilitate CO_2 fixation in cyanobacteria and some proteobacteria. *Microbiol Mol Biol R* **77**, 357-379 (2013).
8. J. M. Shively, F. Ball, D. H. Brown, R. E. Saunders, Functional organelles in prokaryotes: polyhedral inclusions (carboxysomes) of *Thiobacillus neapolitanus*. *Science* **182**, 584-586 (1973).
9. G. D. Price, J. R. Coleman, M. R. Badger, Association of carbonic-anhydrase activity with carboxysomes isolated from the cyanobacterium *Synechococcus* PCC7942. *Plant Physiol* **100**, 784-793 (1992).
10. S. Tanaka, M. R. Sawaya, M. Phillips, T. O. Yeates, Insights from multiple structures of the shell proteins from the beta-carboxysome. *Protein Sci* **18**, 108-120 (2009).
11. F. Cai *et al.*, The structure of CcmP, a tandem bacterial microcompartment domain protein from the beta-carboxysome, forms a subcompartment within a microcompartment. *J Biol Chem* **288**, 16055-16063 (2013).
12. S. Tanaka *et al.*, Atomic-level models of the bacterial carboxysome shell. *Science* **319**, 1083-1086 (2008).

13. T. J. Keeling, B. Samborska, R. W. Demers, M. S. Kimber, Interactions and structural variability of beta-carboxysomal shell protein CcmL. *Photosynth Res* **121**, 125-133 (2014).

14. B. D. Rae, B. M. Long, M. R. Badger, G. D. Price, Structural determinants of the outer shell of beta-carboxysomes in *Synechococcus elongatus* PCC 7942: roles for CcmK2, K3-K4, CcmO, and CcmL. *PLoS One* **7**, e43871 (2012).

15. B. M. Long, M. R. Badger, S. M. Whitney, G. D. Price, Analysis of carboxysomes from Synechococcus PCC7942 reveals multiple Rubisco complexes with carboxysomal proteins CcmM and CcaA. *J Biol Chem* **282**, 29323-29335 (2007).

16. S. S. Cot, A. K. So, G. S. Espie, A multiprotein bicarbonate dehydration complex essential to carboxysome function in cyanobacteria. *J Bacteriol* **190**, 936-945 (2008).

17. B. M. Long, L. Tucker, M. R. Badger, G. D. Price, Functional cyanobacterial beta-Carboxysomes have an absolute requirement for both long and short forms of the CcmM protein. *Plant Physiol* **153**, 285-293 (2010).

18. J. N. Kinney, A. Salmeen, F. Cai, C. A. Kerfeld, Elucidating essential role of conserved carboxysomal protein CcmN reveals common feature of bacterial microcompartment assembly. *J Biol Chem* **287**, 17729-17736 (2012).

19. M. T. Lin *et al.*, beta-Carboxysomal proteins assemble into highly organized structures in *Nicotiana* chloroplasts. *Plant J* **79**, 1-12 (2014).

20. J. C. Cameron, S. C. Wilson, S. L. Bernstein, C. A. Kerfeld, Biogenesis of a bacterial organelle: the carboxysome assembly pathway. *Cell* **155**, 1131-1140 (2013).

21. R. H. Kohler, J. Cao, W. R. Zipfel, W. W. Webb, M. R. Hanson, Exchange of protein molecules through connections between higher plant plastids. *Science* **276**, 2039-2042 (1997).

22. B. Samborska, M. S. Kimber, A dodecameric CcmK2 structure suggests beta-carboxysomal shell facets have a double-layered organization. *Structure* **20**, 1353-1362 (2012).

23. T. Kaneko *et al.*, Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions (supplement). *DNA Res* **3**, 185-209 (1996).

24. C. D. Hu, T. K. Kerppola, Simultaneous visualization of multiple protein interactions in living cells using multicolor fluorescence complementation analysis. *Nat Biotechnol* **21**, 539-545 (2003).

25. M. R. Hanson, M. T. Lin, A. E. Carmo-Silva, M. A. Parry, Towards engineering carboxysomes into C3 plants. *Plant J* **87**, 38-50 (2016).

26. W. Bonacci *et al.*, Modularity of a carbon-fixing protein organelle. *Proc Natl Acad Sci U S A* **109**, 478-483 (2012).

27. N. M. Mangan, A. Flamholz, R. D. Hood, R. Milo, D. F. Savage, pH determines the energetic efficiency of the cyanobacterial CO_2 concentrating mechanism. *Proc Natl Acad Sci U S A* **113**, E5354-5362 (2016).

28. S. C. Huber, Effect of pH on chloroplast photosynthesis. Inhibition of O2 evolution by inorganic phosphate and magnesium. *Biochim Biophys Acta* **545**, 131-140 (1979).

29. M. J. Gething, J. Sambrook, Protein folding in the cell. *Nature* **355**, 33-45 (1992).

30. C. J. Roberts, Non-native protein aggregation kinetics. *Biotechnol Bioeng* **98**, 927-938 (2007).

31. J. Kudla, R. Bock, Lighting the way to protein-protein interactions: recommendations on best practices for bimolecular fluorescence complementation analyses. *Plant Cell* **28**, 1002-1008 (2016).

32. I. A. Sparkes, J. Runions, A. Kearns, C. Hawes, Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. *Nat Protoc* **1**, 2019-2025 (2006).